

## Personalização de Serviços de Comércio Eletrônico

## Personalização

- Adaptação das informações e dos recursos providos aos usuários com a finalidade de incentivar transações comerciais.
- Não afeta a semântica dos serviços, pois aspectos funcionais não são alterados.

## Personalização de Serviços no Café

- Nível Conceitual: inalterado
- Nível de Aplicação: inalterado
- Nível Funcional:
  - Especificação da personalização
  - O que será personalizado?
- Nível de Execução:
  - Detalhamento de todo o processo de personalização

## Nível Funcional

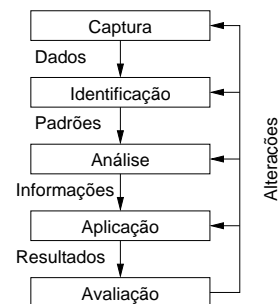
Para cada serviço:

- Finalidade
  - facilitar utilização, promover vendas
- Informações personalizadas
  - quais informações e procedimentos são personalizados
- Critérios de personalização
  - preferências dos clientes, produtos correlacionados

## Nível de Execução

- Fases:
  - Captura: quais os dados a serem coletados
  - Identificação padrões: determina padrões de comportamento individuais ou coletivos
  - Análise: caracteriza o comportamento
  - Aplicação: personaliza serviços
  - Avaliação: verifica eficácia e determina alterações

## Nível de Execução



## Nível de Execução Captura de Dados

Fontes de dados:

- Acessos
  - registro das requisições atendidas pelo servidor.
- Referências
  - páginas (e portanto sítios) a partir dos quais o servidor é acessado
- Agentes
  - requisições feitas através de agentes de busca ou negociação
- Cadastro
  - informações cadastrais providas pelos clientes
- Informações sobre produtos
  - categoria, descrição e composição

## Fonte de Dados Acessos

- Campos:
  - IP
  - Data-Hora
  - Requisição
  - Resultado
  - Bytes Enviados
- IP não é suficiente para identificação!
  - IPs obtidos dinamicamente
  - Proxies

## Representações de Dados

- Expressões regulares
- Hierarquias de conceito
- Regras
- Grafos:
  - CBMG
  - Gramáticas probabilísticas

## Representações de Dados Expressões Regulares

Utilizadas para representar seqüências de requisições de clientes.

- Templates de navegação
- Notação:
  - < - Início de uma seqüência
  - > - Fim de uma seqüência
  - | - Páginas conectadas foram acessadas na mesma visita
  - \* - Zero ou mais acessos a qualquer páginas
  - ; - Páginas conectadas foram acessadas em visitas diferentes
  - ^ - Página referenciada não foi acessada
  - ? - Um acesso a qualquer página

## Representação de Dados Expressões Regulares

**Exemplo:** Análise da eficácia de uma campanha de presentes, focando em consumidores que entraram pela página principal, viram o anúncio da campanha e posteriormente voltaram e compraram. Excluir caminhos que incluam visitas às ofertas padrão da loja.

Exemplo:

1. <index.htm|\*/gifts.htm\*;purchase.htm?>
2. <\*;^ reduced.htm;\*>
3. <\*;^ junk.htm;\*>

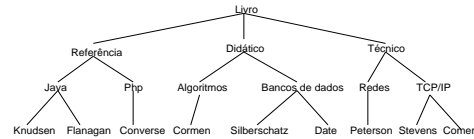
## Representação de Dados Expressões Regulares

- Forma simples e rápida de verificar hipóteses a respeito do comportamento dos usuários
- Determinar o conjunto de expressões que melhor representa uma seqüência de acessos é um problema complexo
- Verificar a correção e a validade de expressões regulares é um problema ainda mais complexo

## Representação de Dados Hierarquias de Conceito

- Organizam informações de acordo com critérios semânticos
- Realizada manualmente
- Critérios aplicados em cada nível têm que ser definidos previamente à construção e a mudança de critérios ou sua ordem demanda a reconstrução
- Exemplo:
  - divisão de livros em gêneros
- Utilizadas para correlacionar produtos
  - permitem correlacionar comportamento de usuários

## Representação de Dados Hierarquias de Conceitos



## Representação de Dados Regras

- Indicam o comportamento probabilístico do usuário
- Exemplos:
  - 40% dos usuários que acessam a página principal vão à página de ofertas
- A representação indica ações de caminhamento, sem entrar no mérito semântico.

## Representação de Dados Regras

Cliente1	Cliente2	Cliente3	Cliente4	Cliente5
Iniciar	Iniciar	Buscar	Iniciar	Iniciar
Buscar	Buscar	Buscar	Navegar	Navegar
Ver	Buscar	Buscar	Navegar	Buscar
Alocar	Ver	Buscar	Ver	Alocar
Pagar	Ver	Buscar	Alocar	Alocar

## Representação de Dados Regras

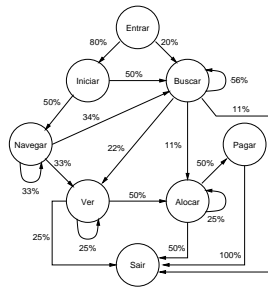
- Regras
  - 80% dos clientes iniciam a sua interação pela página inicial
  - 50% dos clientes que acessam a página inicial realizam buscas
  - 22% das buscas são seguidas por visualização de produtos
  - 25% das alocações resultam em vendas
- Observações
  - Número de regras pode ser muito grande
  - Definir a significância de uma regra depende da finalidade da análise.

## Representação de Dados Grafos

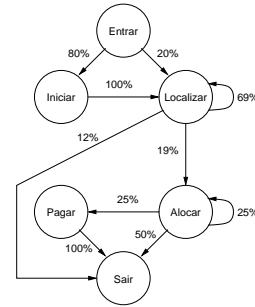
Forma generalizada de representar regras de caminhamento.

- CBMG:
  - Vértices: estados de compra.
  - Arestas: probabilidade de ir para um estado a partir de um estado corrente
- Gramáticas Probabilísticas de Hipertexto:
  - Vértices: páginas
  - Arestas: probabilidade de ir para uma página a partir da página corrente

## Representação de Dados Gramáticas Probabilísticas



## Representação de Dados CBMG



## Representação de Dados CBMG

- Observações
  - Apesar de CBMGs serem mais abstratos, são considerados uma representação de dados, pois o mapeamento entre páginas e entidades é feito de acordo com a especificação dos serviços.
  - As representações baseadas em grafos podem ser utilizadas para identificar perfis de comportamento de usuários e classificá-los em compradores prováveis ou não.

## Identificação de Padrões

Padrões são modelos que representam a interação de um cliente

- Técnicas
  - Inteligência artificial
  - Mineração de dados
  - Psicologia
  - Teoria da Informação
- Exemplos:
  - 40% dos clientes que acessaram A também acessaram B
  - 30% dos clientes que acessaram A compraram um produto
  - 60% dos clientes que compraram X, também compraram Y no prazo de 15 dias
  - 50% dos clientes que compraram C tem entre 20 e 25 anos e moram em BH

## Processo de Identificação de Padrões

- Fases:
  - Limpeza de dados
  - Identificação de transações
  - Integração de dados
  - Transformações
  - Determinação de padrões
  - Análise de padrões

## Identificação de Padrões

- Padrões informativos
  - explicitam informações interessantes
  - agrupamentos, regras de associação
- Padrões de previsão
  - estimativa de atributo desconhecido com base em dados conhecidos
  - regressão, classificação

## Agrupamentos (Clusters)

- Algoritmos identificam características comuns a grupos de entidades de forma a particionar o universo de entidades de acordo com os seus atributos
- As entidades de um dado agrupamento são similares em relação a um dado critério
- Agrupamentos podem ser utilizados como ponto de partida para a exploração de relacionamentos entre entidades.
- Questões:
  - Quais atributos são relevantes?
  - Quais os critérios de similaridade?
- Centróide:
  - ponto cujo valor é a média dos pontos do agrupamento

## Agrupamentos Exemplo

Cada coluna corresponde às ações tomadas por um cliente.  
Cada entrada contém o número de ações realizadas do respectivo tipo.

	$C_1$	$C_2$	$C_3$	$C_4$	$C_5$
Iniciar	1,00	1,00	0,00	1,00	1,00
Localizar	2,00	4,00	5,00	3,00	2,00
Alocar	1,00	0,00	0,00	1,00	2,00
Pagar	1,00	0,00	0,00	0,00	0,00

## Algoritmo K-means

- Escolha k pontos a serem utilizados como estimativa inicial para os centróides
- Repita
  - Examine cada ponto a ser agrupado e o associe ao centróide mais próximo, atualizando o valor do centróide
- Até que nenhum ponto mude de cluster ou o máximo de iterações seja atingido.

## Algoritmo K-Means Exemplo

Primeira iteração.

Inicialização aleatória

$C_2$  é ponto inicial

$C_3$  é ponto inicial

	$A_2$	$A_3$
Iniciar	1,00	0,00
Localizar	4,00	5,00
Alocar	0,00	0,00
Pagar	0,00	0,00

## Algoritmo K-Means Exemplo

Primeira iteração.

Associando o Cliente 1

$$Dist(C_1, A_2) = 4,000$$

$$Dist(C_1, A_3) = 6,000$$

	$A_{12}$	$A_3$
Iniciar	1,00	0,00
Localizar	3,00	5,00
Alocar	0,50	0,00
Pagar	0,50	0,00

## Algoritmo K-Means Exemplo

Primeira iteração.

Associando o Cliente 4

$$Dist(C_4, A_{12}) = 1,000$$

$$Dist(C_4, A_3) = 4,000$$

	$A_{124}$	$A_3$
Iniciar	1,00	0,00
Localizar	3,00	5,00
Alocar	0,67	0,00
Pagar	0,33	0,00

## Algoritmo K-Means Exemplo

Primeira iteração.

Associando o Cliente 5

$$Dist(C_5, A_{124}) = 2,667$$

$$Dist(C_5, A_3) = 6,000$$

	$A_{1245}$	$A_3$
Iniciar	1,00	0,00
Localizar	2,75	5,00
Alocar	1,00	0,00
Pagar	0,25	0,00

## Algoritmo K-Means Exemplo

Segunda iteração.

Associando o Cliente 1

$$Dist(C_1, A_{1245}) = 1,500$$

$$Dist(C_1, A_3) = 6,000$$

	$A_{1245}$	$A_3$
Iniciar	1,00	0,00
Localizar	2,75	5,00
Alocar	1,00	0,00
Pagar	0,25	0,00

## Algoritmo K-Means Exemplo

Segunda iteração.

Associando o Cliente 2

$$Dist(C_2, C_{1245}) = 2,500$$

$$Dist(C_2, C_3) = 2,000$$

	$A_{145}$	$A_{23}$
Iniciar	1,00	0,50
Localizar	2,75	4,50
Alocar	1,00	0,00
Pagar	0,25	0,00

## Algoritmo K-Means Exemplo

Segunda iteração.

Associando o Cliente 3

$$Dist(C_3, A_{145}) = 4,500$$

$$Dist(C_3, A_{23}) = 1,000$$

	$A_{145}$	$A_{23}$
Iniciar	1,00	0,50
Localizar	2,75	4,50
Alocar	1,00	0,00
Pagar	0,25	0,00

## Algoritmo K-Means Exemplo

Segunda iteração.

Associando o Cliente 4

$$Dist(C_4, A_{145}) = 0,500$$

$$Dist(C_4, A_{23}) = 3,000$$

	$A_{145}$	$A_{23}$
Iniciar	1,00	0,50
Localizar	2,33	4,50
Alocar	1,33	0,00
Pagar	0,33	0,00

## Algoritmo K-Means Exemplo

Segunda iteração.

Associando o Cliente 5

$$Dist(C_5, A_{145}) = 1,333$$

$$Dist(C_5, A_{23}) = 5,000$$

	$A_{145}$	$A_{23}$
Iniciar	1,00	0,50
Localizar	2,33	4,50
Alocar	1,33	0,00
Pagar	0,33	0,00

## Algoritmo K-means

Questões

- Função similaridade
  - distância geométrica
  - outras medidas de distância
  - proximidade ortográfica
  - proximidade semântica
- Número de clusters
  - função da finalidade e da natureza dos pontos

## Algoritmo Árvore Geradora Mínima

1. Faça cada ponto um cluster
2. Enquanto numcluster > objetivo
  - Determine o centróide de cada cluster
  - Calcule a matriz de similaridade entre os clusters
  - Determine os clusters com maior similaridade não máxima (ou seja, idênticos) e faça um novo cluster

## Árvore Geradora Mínima Exemplo

Agrupamentos

	$A_1$	$A_2$	$A_3$	$A_4$	$A_5$
Iniciar	1,00	1,00	0,00	1,00	1,00
Localizar	2,00	4,00	5,00	3,00	2,00
Alocar	1,00	0,00	0,00	1,00	2,00
Pagar	1,00	0,00	0,00	0,00	0,00

## Árvore Geradora Mínima Exemplo

Distâncias

	$A_1$	$A_2$	$A_3$	$A_4$	$A_5$
$A_1$	0,00	4,00	6,00	2,00	2,00
$A_2$	4,00	0,00	2,00	2,00	4,00
$A_3$	6,00	2,00	0,00	4,00	6,00
$A_4$	2,00	2,00	4,00	0,00	2,00
$A_5$	2,00	4,00	6,00	2,00	0,00

## Árvore Geradora Mínima Exemplo

Agrupamentos				
	$A_2$	$A_3$	$A_5$	$A_{14}$
Iniciar	1,00	0,00	1,00	1,00
Localizar	4,00	5,00	2,00	2,50
Alocar	0,00	0,00	2,00	1,00
Pagar	0,00	0,00	0,00	0,50

Distâncias				
	$A_2$	$A_3$	$A_5$	$A_{14}$
$A_2$	0,00	2,00	4,00	3,00
$A_3$	2,00	0,00	6,00	5,00
$A_5$	4,00	6,00	0,00	2,00
$A_{14}$	3,00	5,00	2,00	0,00

## Árvore Geradora Mínima Exemplo

Agrupamentos				Distâncias			
	$A_5$	$A_{14}$	$A_{23}$		$A_5$	$A_{14}$	$A_{23}$
Iniciar	1,00	1,00	0,50	$A_5$	0,00	2,00	5,00
Localizar	2,00	2,50	4,50	$A_{14}$	2,00	0,00	4,00
Alocar	2,00	1,00	0,00	$A_{23}$	5,00	4,00	0,00
Pagar	0,00	0,50	0,00				

## Árvore Geradora Mínima Exemplo

Agrupamentos			Distâncias		
	$A_{23}$	$A_{145}$		$A_{23}$	$A_{145}$
Iniciar	0,50	1,00	$A_{23}$	0,00	4,33
Localizar	4,50	2,33	$A_{145}$	4,33	0,00
Alocar	0,00	1,33			
Pagar	0,00	0,33			

## Algoritmo Árvore Geradora Mínima

### Questões

- Número de clusters ainda é um problema
  - podemos usar distância mínima como critério de parada?
- Função similaridade
  - Qual é?
  - Qual o seu significado?

## Regras de Associação

- padrões informativos na forma  $X \rightarrow Y$ , onde  $X$  e  $Y$  são conjuntos disjuntos de instâncias dos atributos de uma entrada.
- Métricas:
  - Suporte: probabilidade que um registro contenha tanto  $X$  quanto  $Y$
  - Confiança: probabilidade que um registro contenha  $Y$ , dado que contém  $X$
- Exemplo: análise de dados de supermercados  
50% das pessoas que compram  $A$  e  $B$ , compram  $C$

## Geração de Regras de Associação

- Problema: Gerar todas as regras que tenham suporte e confiança maiores que os valores mínimos especificados (*minsup* e *minconf*, respectivamente).
- Duas etapas:
  1. Determinar todos os conjuntos de itens cujo suporte é maior do que o suporte mínimo especificado. Eles são chamados de conjuntos frequentes de itens.
  2. Gerar as regras de associação a partir dos conjuntos frequentes de itens. Dados conjuntos frequentes  $XYZ$  e  $XY$ , a regra  $XY \rightarrow Z$  é válida se

$$\frac{\text{sup}(XYZ)}{\text{sup}(XY)} > \text{minconf}$$

## O Algoritmo Apriori Notação:

- Seja  $k$  – *itemset* um conjunto de itens de tamanho  $k$ .
- Seja  $F_k$  o conjunto frequente de itens de tamanho  $k$ .
- Seja  $C_k$  o conjunto de itens candidatos (potencialmente frequentes) de tamanho  $k$ .

## O Algoritmo Apriori

```

1   $F_1 = \{\text{itens frequentes}\}$ 
2  for ( $k=2$ ;  $|F_{k-1}| > 0$ ;  $k++$ )
3       $C_k = \text{gera\_candidatos}(F_{k-1})$ 
4       $\forall$  transações  $T \in DB$ 
5           $\forall$  subconjuntos  $t \in T$ 
6              if ( $c \in C_k$  e  $c \subseteq t$ )
7                  c.contador ++
8       $F_k = \{c \in C_k \text{ e } c.\text{contador} \geq \text{minsup}\}$ 
9       $\forall F_k, k > 2$ 
10     gera_regras( $F_k$ )
```

## Geração de Conjuntos Candidatos

```

1  gera_candidatos( $F_{k-1}$ )
2     $C_k = \emptyset$ 
3     $\forall p, q | p, q \in F_{k-1}$ 
4      if ( $p_i == q_i, 1 \leq i \leq k-2$  e  $p_{k-1} < q_{k-1}$ )
5         $C_k = C_k \cup (p \cup q)$ 
6     $\forall c \in C_k$ 
7       $\forall s | s \subset c$  e  $|s| = k-1$ 
8        if ( $s \notin F_{k-1}$ )
9           $C_k = C_k - c$ 
10   return  $C_k$ 

```

## Geração de Conjuntos Candidatos Exemplo:

### • Situação:

$$F_3 = \{\{1, 2, 3\}, \{1, 2, 4\}, \{1, 3, 4\}, \{1, 3, 5\}, \{2, 3, 4\}, \{3, 4, 5\}\}$$

### • Enumeração:

$$C_4 = \{\{1, 2, 3, 4\}, \{1, 3, 4, 5\}\}$$

### • Corte:

$$C_4 = \{\{1, 2, 3, 4\}\}$$

pois  $\{1, 4, 5\}$  não é freqüente.

## Apriori Exemplo

Transações a serem mineradas. Suporte: 20%. Confiança: 70%.

```

1  A B C
2  A B
3  A B C
4  C E
5  A B C D
6  C D E
7  B C D
8  B C E
9  C D
10 A C D

```

## Apriori Exemplo

$F_1$		$C_2$			$F_2$
Conj.	Sup.	Enum.	Corte	Sup.	Conj.
A	5 (50%)	AB	AB	4 (40%)	AB
B	6 (60%)	AC	AC	4 (40%)	AC
C	9 (90%)	AD	AD	2 (20%)	AD
D	5 (50%)	AE	AE	0 (00%)	BC
E	3 (30%)	BC	BC	5 (50%)	BD
		BD	BD	2 (20%)	CD
		BE	BE	1 (10%)	CE
		CD	CD	5 (50%)	
		CE	CE	3 (30%)	
		DE	DE	1 (10%)	

## Apriori Exemplo

$F_2$		$C_3$			$F_3$
Conj.	Sup.	Enum.	Corte	Sup.	Conj.
AB	4 (40%)	ABC	ABC	3 (30%)	ABC
AC	4 (40%)	ABD	ABD	1 (10%)	ACD
AD	2 (20%)	ACD	ACD	2 (20%)	BCD
BC	5 (50%)	BCD	BCD	2 (20%)	
BD	2 (20%)	CDE			
CD	5 (50%)				
CE	3 (30%)				

## Geração de Regras

```

1  gera_regras ( $F_k$ )
2     $R_k = \emptyset$ 
3     $\forall s | s \in F_k$ 
4       $\forall a | a \subset s$ 
5         $c = s - a$ 
6         $conf = \frac{sup_a}{sup_s}$ 
7        if ( $conf > minconf$ )
8           $R_k = R_k \cup (a \rightarrow c)$ 
9   return  $R_k$ 

```

## Apriori Exemplo

Geração das regras referentes a  $ABC$

Regras	Confiança
$AB \rightarrow C$	75%
$AC \rightarrow B$	75%
$BC \rightarrow A$	60%
$A \rightarrow BC$	60%
$B \rightarrow AC$	50%
$C \rightarrow AB$	33%

## Apriori Exemplo

Conj. Base		Regra de Associação	
Conj.	Sup.	Regra	Confiança
ABC	3 (30%)	$AB \rightarrow C$	75%
		$AC \rightarrow B$	75%
ACD	2 (20%)	$AD \rightarrow C$	100%
BCD	2 (20%)	$BD \rightarrow C$	100%
AB	4 (40%)	$A \rightarrow B$	80%
AC	4 (40%)	$A \rightarrow C$	80%
BC	5 (50%)	$B \rightarrow C$	83%
CD	5 (50%)	$D \rightarrow C$	100%
CE	3 (30%)	$E \rightarrow C$	100%

## Regressão

- Técnicas de descoberta de padrões de previsão onde os padrões a serem previstos são números reais
- Técnicas
  - regressão linear
  - redes neurais
  - CART (*Classification and Regression Tree*)

## Classificação

- Tarefa de dividir conjuntos de objetos, entidades ou pessoas em classes
- Critérios
  - valores literais de atributos, como sexo e idade
  - perfis dos elementos das classes
- Desafio: determinar os critérios a serem utilizados
- Escolha de técnicas é função de
  - critérios de classificação
  - natureza da entrada
  - finalidade da classificação

## Redes Neurais

São modelos matemáticos que implementam mapeamentos não-lineares entre conjuntos de entradas e saídas.

- Denominação vem da sua inspiração biológica
- Redes neurais são compostas de um grande número de neurônios interconectados segundo um padrão pré-definido
- Cada neurônio é uma entidade de processamento independente, que recebe um conjunto de entradas e gera uma saída
- Informação é armazenada nos seus pesos e na própria topologia.

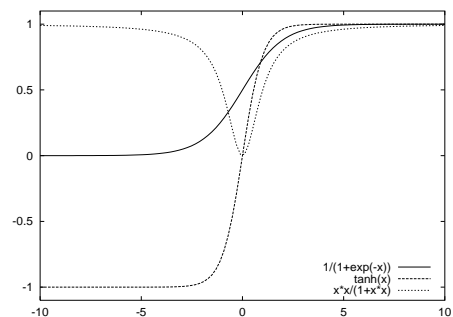
## Redes Neurais

- Uso em classificação:
  - aproximam mapeamentos não lineares entre conjuntos de entradas e saídas
- Exemplo: obtenção de financiamento
  - Entrada: dados cadastrais
  - Saída: taxa de juros e limite de crédito

## Redes Neurais Artificiais Elementos de Processamento

- Vetor de Entradas:  $E = (E_1, \dots, E_i, \dots, E_n)$
  - Vetor de Pesos:  $W_j = (W_{1j}, \dots, W_{ij}, \dots, W_{nj})$
  - Nível de Ativação:  $T_j$
  - Função de Ativação:  $f$
  - Saída:  $S_j$
- $$S_j = f(E * W_j - T_j) = f(\sum_{i=1}^n E_i * W_{ij} - T_j)$$

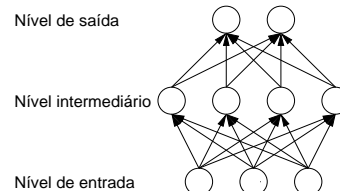
## Funções Sigmóide



## Topologia

- Tipos de conexão:
  - Excitatória
  - Inibitória
- Esquema de interconexão
  - Intra-nível
  - Inter-nível
  - Recorrente

## Redes Neurais Topologia Multinível





## Aprendizagem

- Qualquer mudança no vetor de pesos
- Tipos:
  - Supervisionado: exemplos externos
- Espacial: codifica mapeamentos
- Temporal: codifica sequência de passos
  - Não supervisionado:
- auto-organiza dados, descobrindo propriedades coletivas

## Back-propagation

- Rede feed-forward
  - Não há ciclos
  - Uso para tarefas hetero-associativas
  - Rápida utilização
  - Aprendizagem multi-nível
- demanda a derivação de uma relação entre a correção no nível de saída e os níveis internos
- funções de ativação lineares não adicionam graus de liberdade para redes multi-nível

## Back-propagation: funcionamento

- Aprendizagem
  - Pesos associados às conexões são calculados
  - Baseado em um conjunto de treinamento, contendo valores de entrada e saída.
  - A entrada de cada mapeamento do conjunto de treinamento é submetida à rede, que gera uma saída. A saída gerada é comparada à saída do mapeamento, e a sua diferença usada para ajustar os pesos da rede neuronal.
  - O processo se repete até que o erro entre saídas obtidas e desejadas esteja dentro do limite pré-definido.
- Produção
  - Entradas são submetidas à rede neuronal que gera as saídas de acordo com os padrões treinados.

## Back-propagation

- Questões:
  - parâmetros de treinamento e tamanho da rede devem ser suficientes para comportar o mapeamento desejado
  - natureza de cada entrada e saída deve ser compatível com a capacidade e funcionamento da rede
  - conjunto de treinamento deve garantir cobertura e justiça

## Redes Neurais Exemplo

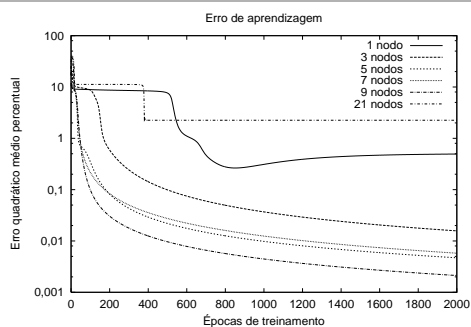
Perfis de Clientes

	$C_1$	$C_2$	$C_3$	$C_4$	$C_5$
Idade	35	16	?	23	20
Sexo	Masc.	Fem.	?	Masc.	Fem.
Renda	2500	200	?	100	5000
Instrução	Univ.	Fund.	?	Univ.	Médio
Taxa de conversão	100%	20%	?	40%	70%

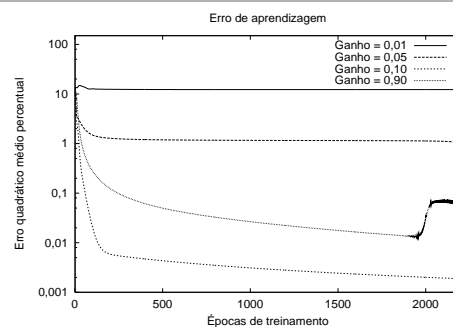
## Redes Neurais Conjunto de treinamento

	$C_1$	$C_2$	$C_4$	$C_5$
Idade	0,50	0,12	0,26	0,20
Sexo	0,00	1,00	0,00	1,00
Renda	0,50	0,04	0,02	1,00
Instrução	1,00	0,60	1,00	0,80
Taxa de conversão	1,00	0,20	0,40	0,70

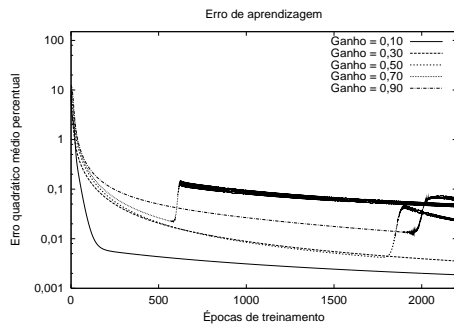
## Redes Neurais Dimensionamento



## Redes Neurais Avaliação da taxa de aprendizagem



## Redes Neurais Avaliação da taxa de aprendizagem



## Análise

- Extrair informações dos padrões de comportamento dos clientes
  - Produtos:
    - baseada em aquisições e dados cadastrais
  - Procedimentos:
    - formas prediletas de interação
  - Formato:
    - formatos e informações apresentadas

## Análise Determinando produtos correlatos

- Ponto de partida:
  - regras de associação
- Tarefa:
  - regras geradas devem ser analisadas quanto à sua significância e aplicabilidade
- Critérios
  - Completude do conjunto de regras - pois pode não haver regras para um dado antecedente
  - Cardinalidade da regras - depende da estratégia de personalização

## Aplicação

- Desenvolver e implementar uma estratégia de personalização a partir da caracterização
- Fatores:
  - mecanismos de personalização
  - escopo
  - forma de atualização

## Mecanismos de Personalização

- Produtos:
  - Venda cruzada: itens complementares
  - Venda ampliada: itens melhores
  - Recomendação: itens de interesse
  - Promoção: itens correlatos em promoção

## Mecanismos de Personalização

- Procedimentos:
  - todos os serviços que recebem parâmetros
  - armazena, para cada usuário, opções e parâmetros utilizados
- Formato:
  - estética: disposição de texto e gráficos
  - conteúdo: tipos de informação a serem apresentados.

## Escopo

- Amplitude do conjunto de usuários a ser personalizado
  - Global: se aplica a todos os clientes ou grupos de clientes
  - Local: se aplica a apenas um cliente
- Compromisso:
  - custo de escopos locais × personalização insipiente

## Atualização

- Especifica a recentidade dos dados usados para personalização:
  - Discreta: atualização periódica
  - Contínua: atualização instantânea

## Aplicação Sugestão de produtos correlatos

- Decisões de implementação para armazenar regras
  - Forma de armazenamento: SGBD × arquivos
  - Critério de seleção das regras
  - Estratégia diante de ausência de informação

## Aplicação Sugestão de produtos correlatos

### Seleção das regras

- Ponto de partida
  - conjunto de seleção – produtos selecionados ou adquiridos pelo cliente
- Estratégia
  - selecionar todas as regras que tenham o conjunto de seleção como antecedente
  - se necessário, selecionar as regras que tenham sobreposição com o antecedente
- Ordenação das regras
  - Porcentagem de cobertura do antecedente
  - Suporte e confiança

## Avaliação

- Métricas:
  - taxa de cliques e taxa de conversão
  - taxa de acerto das informações sugeridas
- Micro-métricas:
  - avaliam a efetividade do sítio
  - visita para listagem
  - listagem para cesta
  - cesta para compra

## Privacidade

- Personalização implica invasão de privacidade
  - Insegurança de usuários depende da informação utilizada
- Política de privacidade - P3P
  - dados coletados
  - utilização
  - controle de acesso
  - anonimidade

## Jambreiro Busca e Navegação

- Facilitar a localização de produtos
- Funcional:
  - destaca produtos de acordo com os interesses dos clientes, os quais são detectados pelos seus dados cadastrais ou operações de localização e compra anteriores.

## Jambreiro Busca e Navegação

### Execução

- Captura de dados:
  - registros de acesso, dados cadastrais e compras anteriores
- Identificação de padrões:
  - segmentar base de produtos: agrupamento
  - correlacionar classes de produtos a dados cadastrais e caminhamento: redes neurais
- Análise:
  - determinar um peso para cada grupo de produtos

## Jambreiro Busca e Navegação

### Execução:

- Aplicação:
  - produtos são enquadrados nas classes
  - classes são ordenadas de acordo com pesos (podem também ser destacados)
- Avaliação:
  - baseada na efetividade das sugestões, isto é, se a sugestão foi aceita

## Jambreiro Colocação na Cesta

- Ampliar a intenção de compra
- Funcional:
  - oferecer produtos correlatos com base em transações anteriores envolvendo o produto sendo colocado na cesta

## Jambreiro Colocação na Cesta

### Execução:

- Captura de dados:
  - escolhas de usuários em uma mesma sessão ou sessões diferentes, conforme o registro de acessos
- Identificação de padrões:
  - determinar correlações entre produtos, mais especificamente regras de associação

## Jambreiro Colocação na Cesta

### Execução:

- Análise:
  - determinar quais regras devem ser usadas para personalização
  - significância das regras: maior suporte
  - fonte: compras, seleções
- Aplicação:
  - selecionar produtos para apresentação, com base nas regras mais similares ao caminharmento corrente do cliente

## Jambreiro Colocação na Cesta

### Execução:

- Avaliação:
  - baseada em eficácia - quais sugestões foram efetivamente adquiridas, mesmo que em uma visita posterior
  - base de regras pode mudar ao longo do tempo, como consequência do comportamento dos clientes também mudar

## Jambreiro Compra

- Facilitar a interação do cliente com o servidor, minimizando a quantidade de informações a ser provida.
- Funcional:
  - agilizar o processo de compra pelo pré-preenchimento das opções e dados de frete e pagamento.

## Jambreiro Compra

### Execução:

- Captura de dados:
  - formulários preenchidos anteriormente
- Identificação de padrões:
  - padrões são simples - valores de campos
- Análise:
  - selecionar o padrão a ser utilizado, tarefa trivial se houver apenas um

## Jambreiro Compra

### Execução:

- Aplicação:
  - simples, pré-preenchimento de campos
- Avaliação:
  - baseada em quantas sugestões são mantidas