

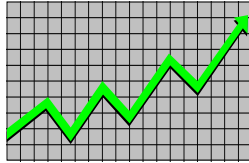
# **Aplicações de Sistemas Inteligentes**

**Germano C. Vasconcelos  
Centro de Informática - UFPE**

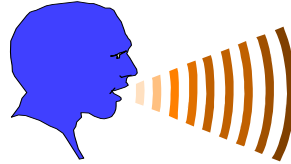
# Roteiro

- Inteligência Computacional
- Aplicações
- Mineração de Dados
- Métodos
- Exemplo em Análise de Crédito

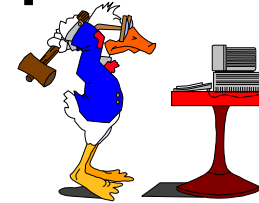
# Aplicações da Inteligência Computacional



Análise de  
mercado



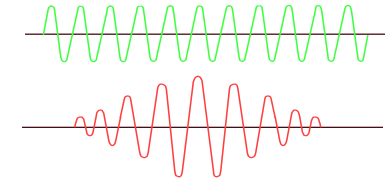
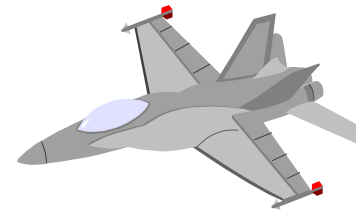
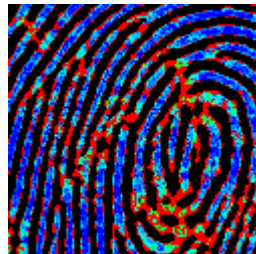
Proc. voz



*mineração de dados*



Análise de  
crédito



Proc. sinais

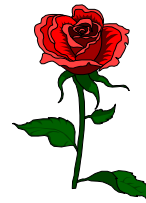


Previsão séries

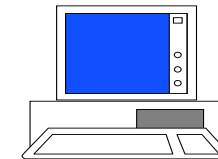
*Luciana de Sales Maciel*



Det. fraudes



Rec. odores

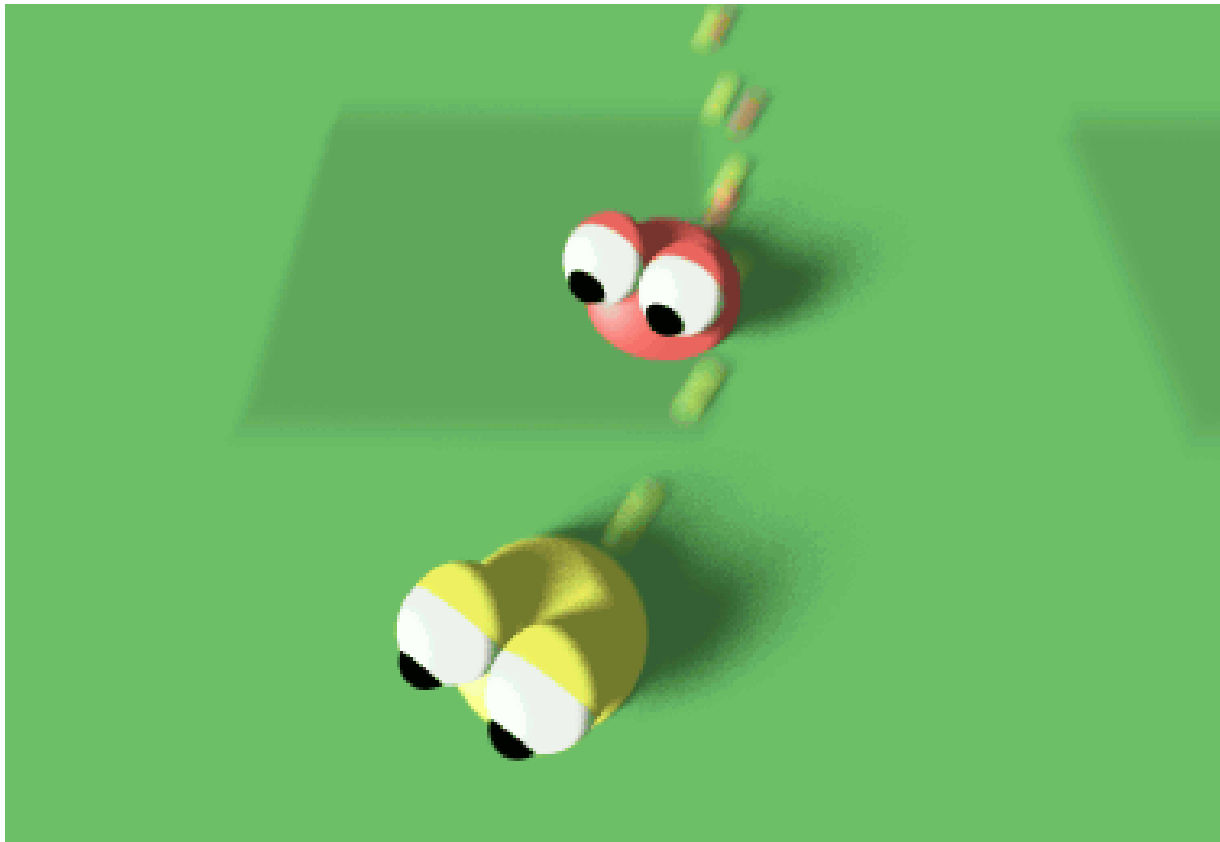


Interfaces

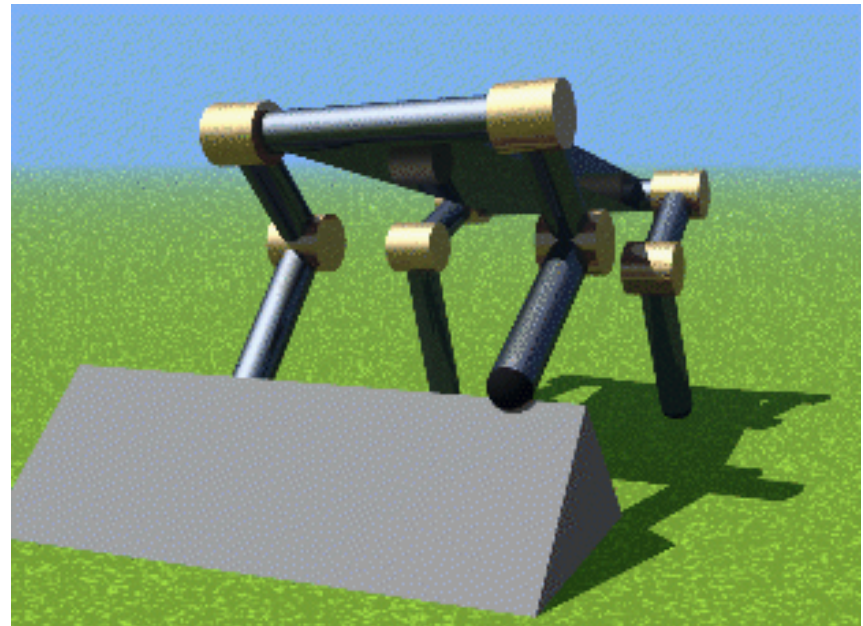
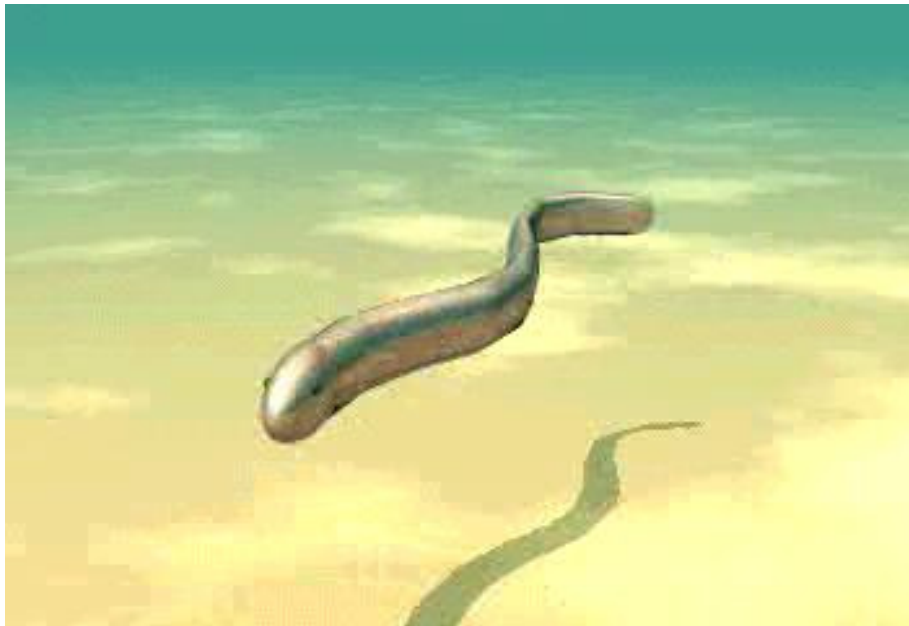


Diagnose  
médica

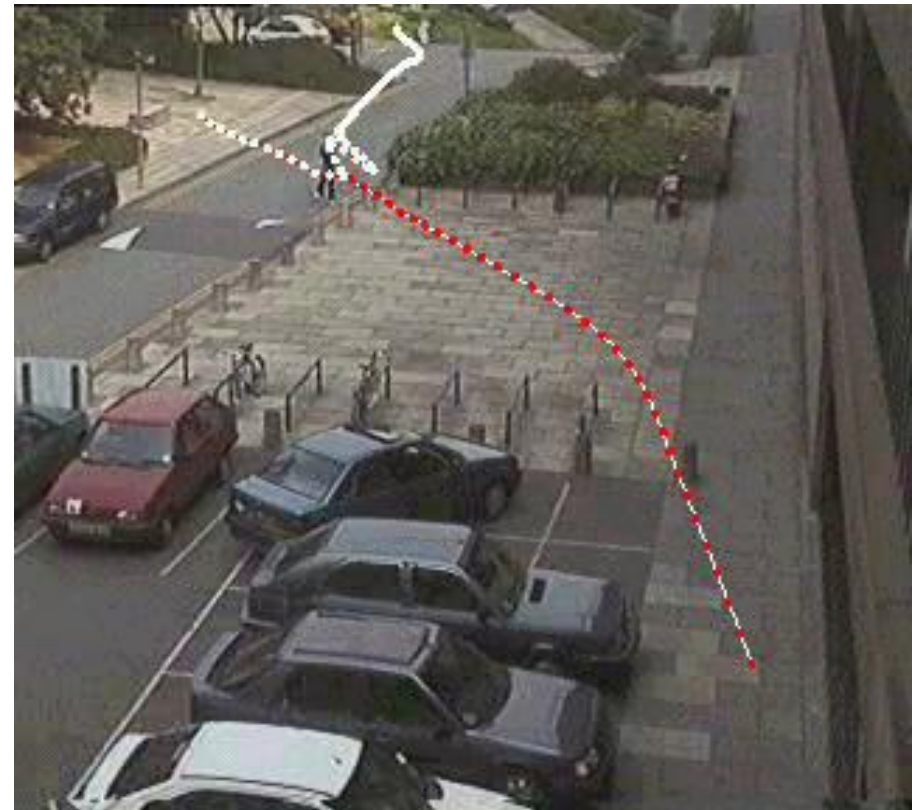
# Organismos Evolutivos: Controle de Perseguição e Fuga



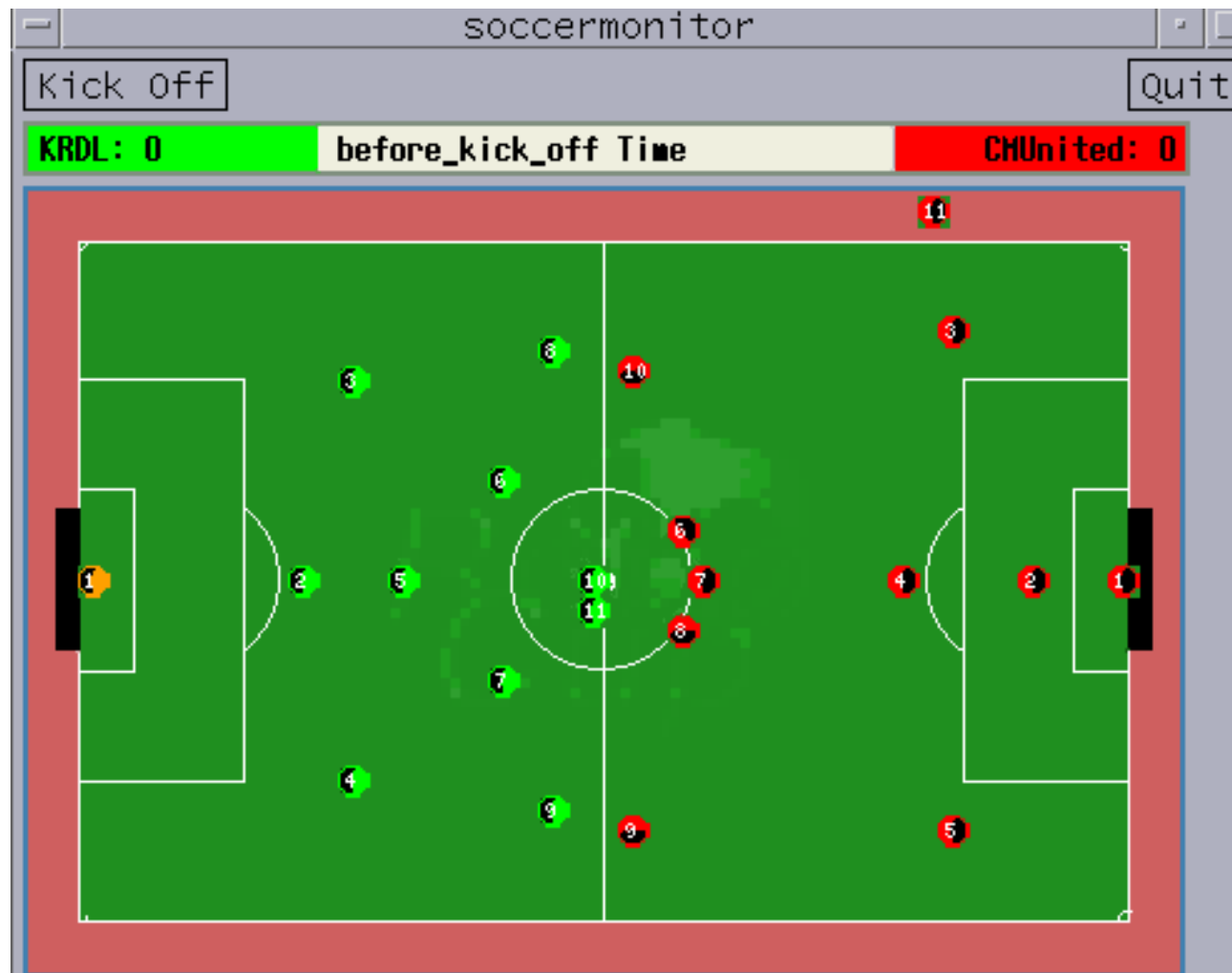
# Vida Artificial



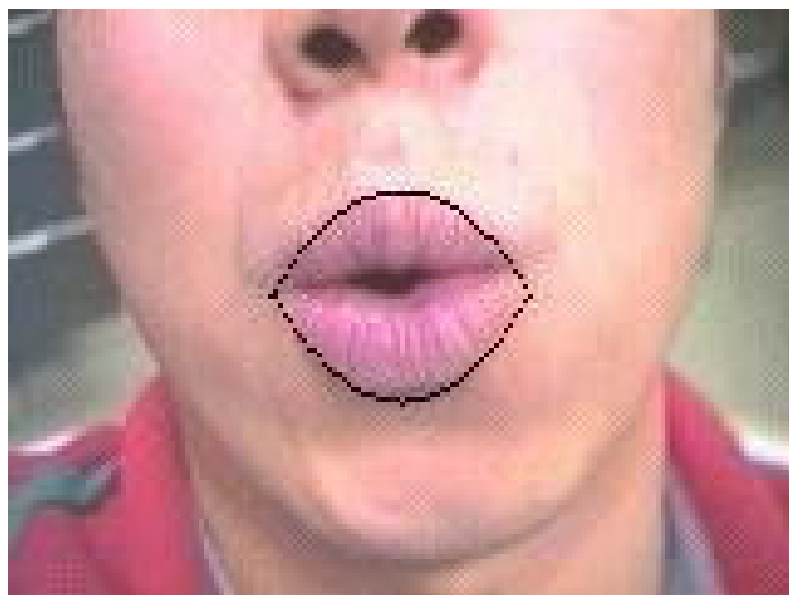
# Estimação de Trajetórias para Reconhecimento de Eventos



# Robot World Cup

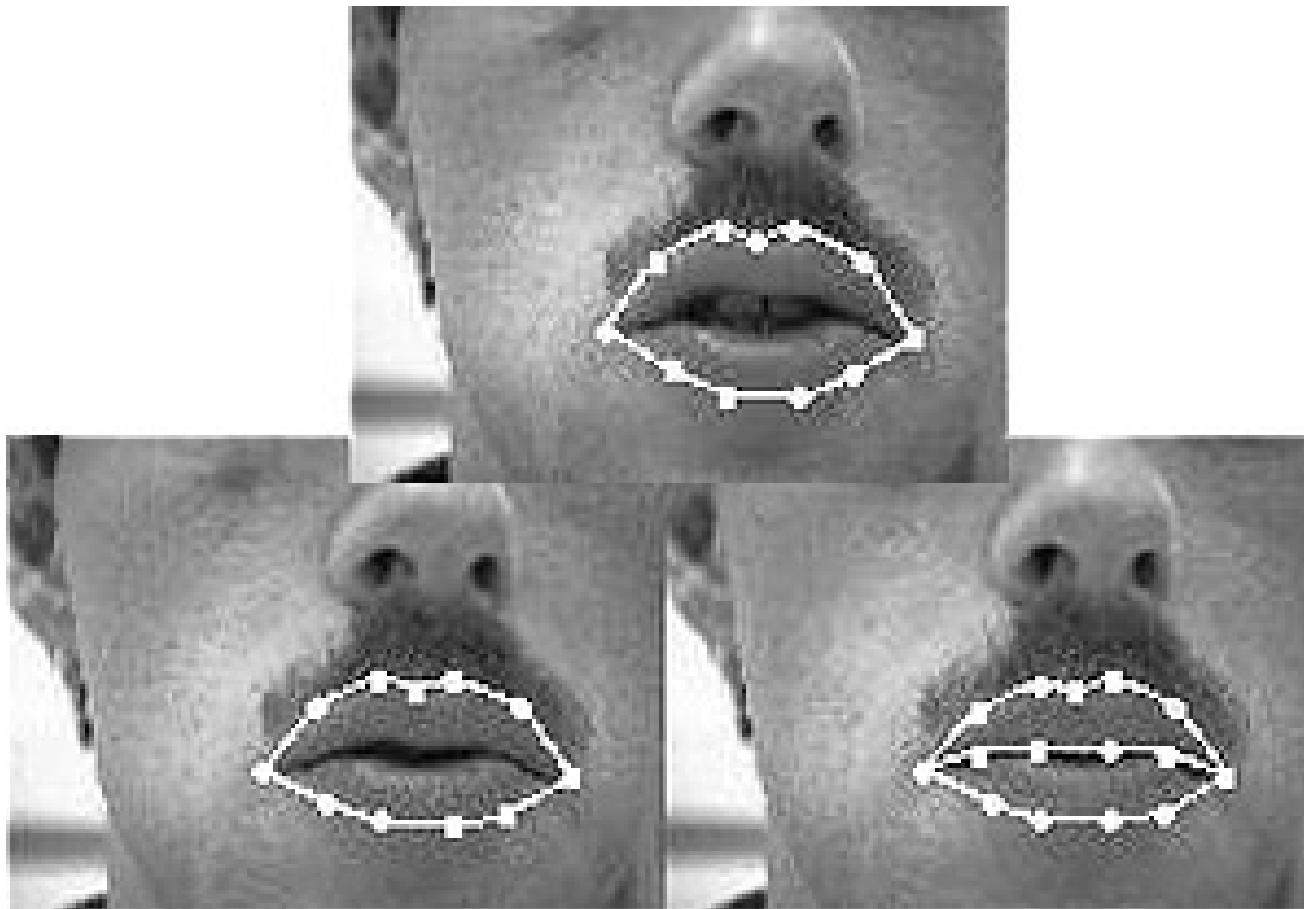


# Conversão de Sons para Imagens

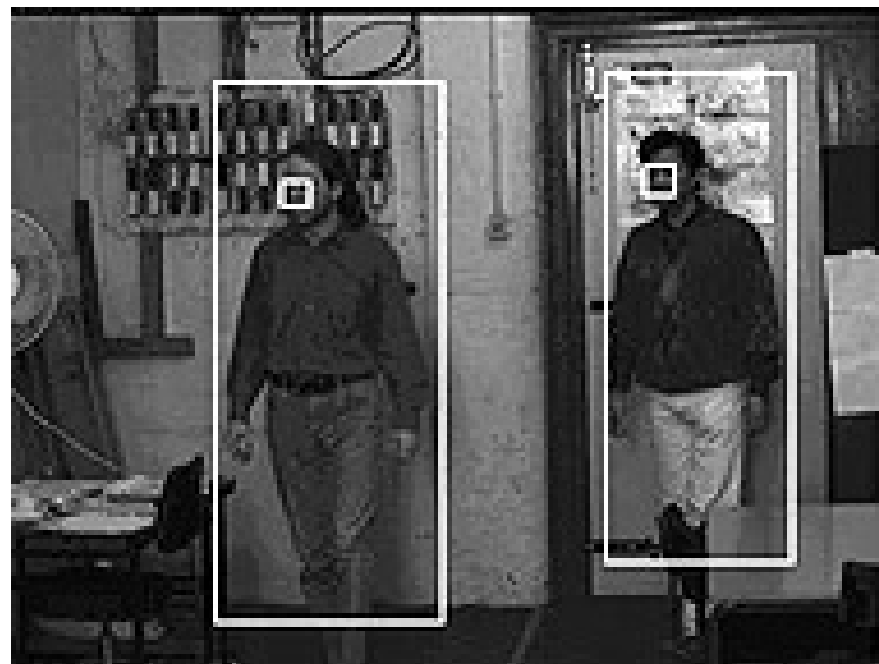




# Leitura Labial



# Monitoramento e Segurança



# Animação Facial



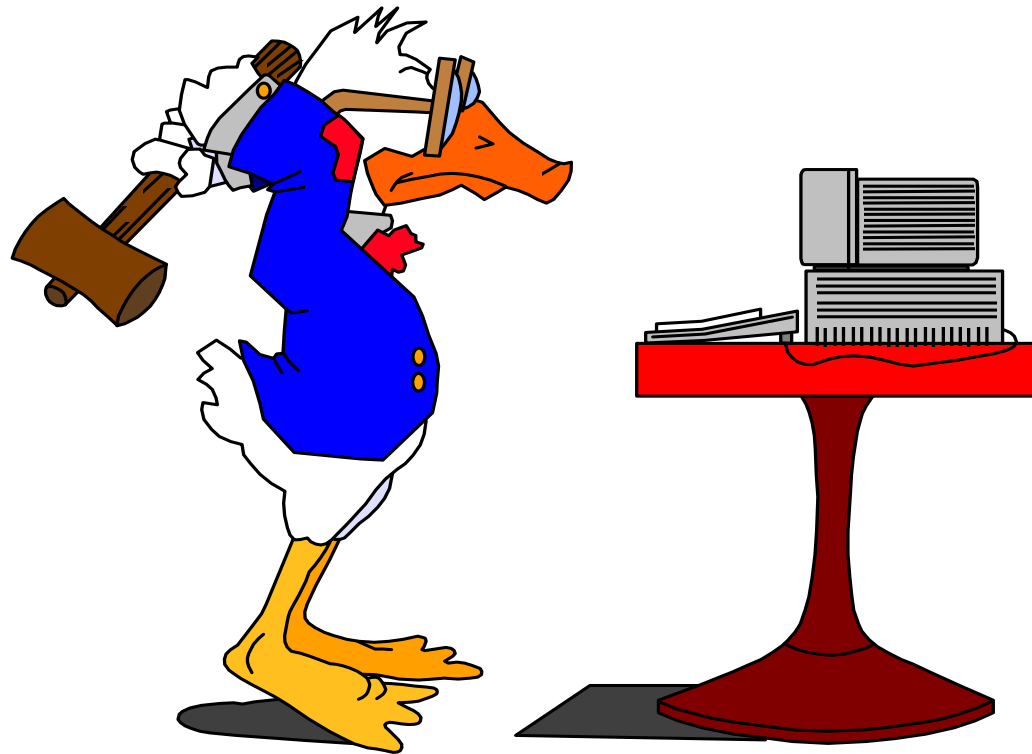
# Animação



# Realidade Virtual



# Mineração de Dados ou *Data Mining*

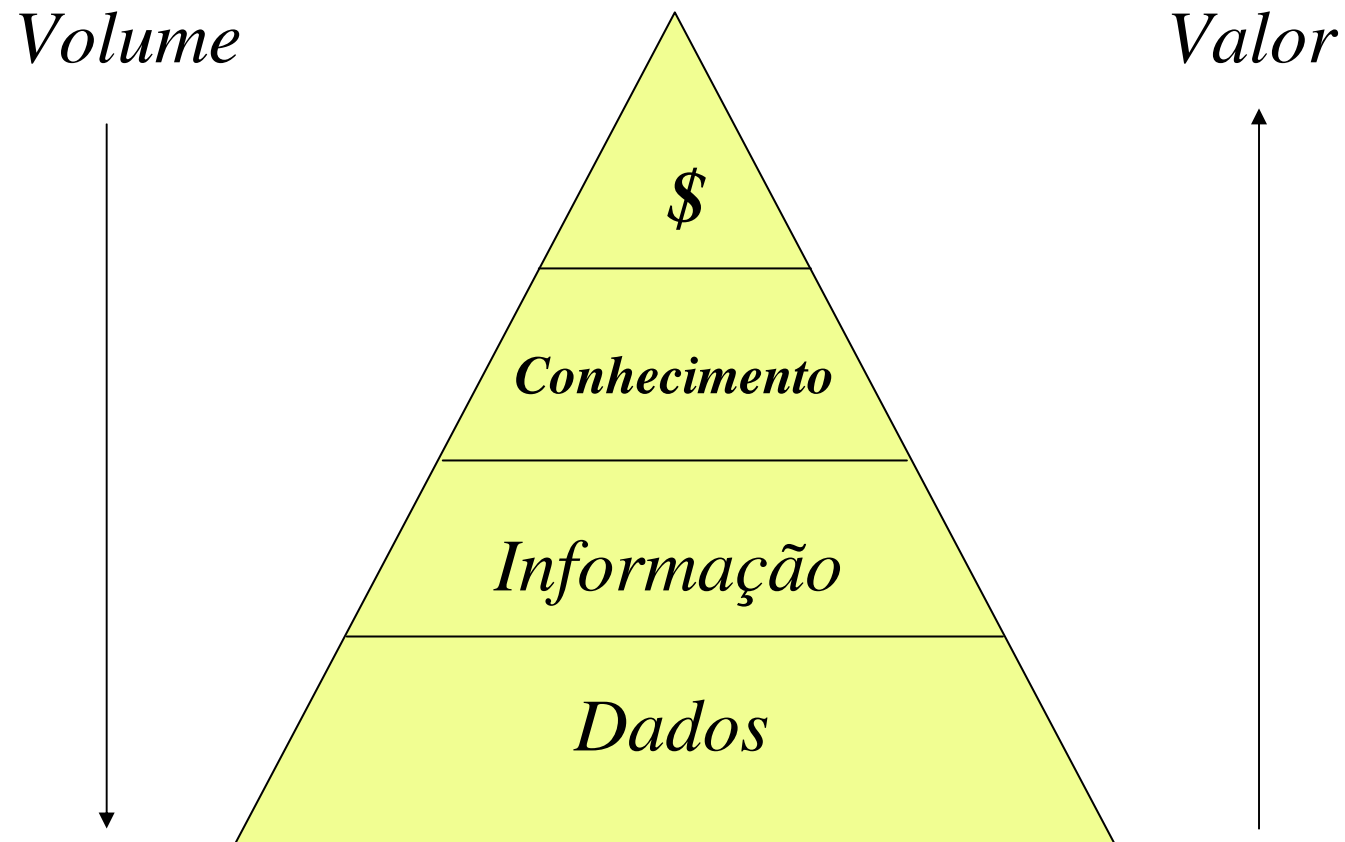


*mineração de dados*

# Mineração de Dados

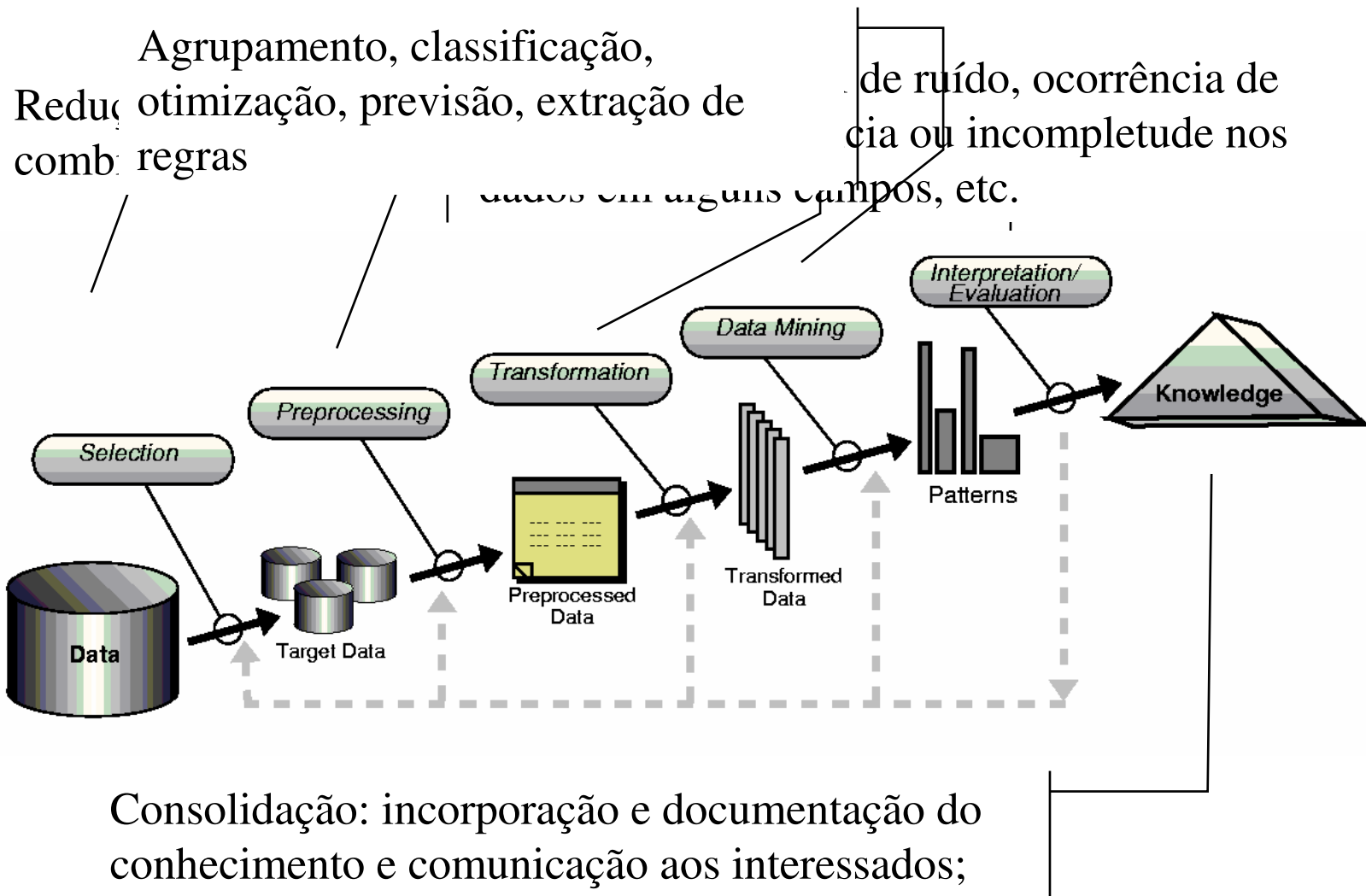
- A mineração de dados objetiva a **transformação de dados em conhecimento útil e estratégico** para a tomada de decisões ...

# Mineração de Dados





# Processo de Descoberta de Conhecimento



# Características Desejáveis do Conhecimento a ser Descoberto

- Correto, o máximo possível...
- Interessante, novo e útil...
- De fácil compreensão para usuários humanos.

# Tarefas Básicas de Mineração de Dados

- Agrupamento
  - Identificação de grupos de indivíduos/registros que têm perfis semelhantes
- Regressão
  - Estimação de valores contínuos na resposta do sistema
- Classificação
  - Decisão do sistema com resposta no domínio discreto
- Extração de regras/descrição
  - Apresentação de relações entre as variáveis e as respostas do sistema

# Consultas a Volumes de Dados : Um exemplo

- Situação
  - banco de dados de vendas de produtos eletrônicos com dados de clientes e produtos
- Consulta direta ao banco de dados (extração de informação)
  - quantas câmeras fotográficas Canon EOS X foram vendidas para o cliente XYZ em agosto de 2001?
- Nível de aplicação
  - atividades do dia a dia da empresa (baixo nível de administração)

# Mineração de Dados : Um exemplo

- Situação
  - banco de dados de vendas de produtos eletrônicos com dados de clientes e produtos
- Conhecimento extraído dos dados
  - Se (idade < 18) e (profissão="estudante") então (compra="camera") (90%)
- Uso do conhecimento descoberto
  - quais os clientes com alta probabilidade de comprar cameras?
- Nível de aplicação (ESTRATÉGICO)
  - mala direta dirigida (database marketing)
  - planejamento de estoque e recursos
  - estabelecimento de políticas de mercado

# Mineração de Dados Possibilita Responder às Seguintes Perguntas:

- Que acontecerá com as vendas do produto A se faltar o produto B nas prateleiras da loja?
- Quanto e quando devo repor o estoque do item A de forma a minimizar o capital imobilizado em estoque ?
- Quanto valerão as ações da empresa X daqui a 15 dias?
- Que produtos devo oferecer em mala direta aos clientes que têm seguro do tipo X, carro do tipo Y e adoram viajar com a família todo ano?
- A quais clientes devo conceder crédito e como definir seus limites?

# Mineração de Dados Possibilita Responder às Seguintes Perguntas:

- Quais os melhores nichos de mercado deste negócio da minha empresa?
- Que segmento do meu mercado devo esperar ter maior retorno em uma campanha?
- Qual o perfil dos piores clientes da minha carteira de crédito?
- 
- Qual o valor ótimo do parâmetro  $k$  para minimizar a perda financeira neste negócio da empresa?
- Que perfis de clientes estão prestes a deixar o nosso serviço pelo da concorrência? Quando isso ocorrerá ?

# Mineração de Dados Possibilita Responder às Seguintes Perguntas:

- Quais as características dos clientes que tendem a cometer o tipo de fraude X?
- Como prevenir potenciais fraudes?
- O que diferencia os clientes que se tornam inadimplentes por falta de capacidade de pagamento daqueles que deixam de pagar por fraude?
- Que tipo de ação de cobrança devo adotar com clientes com certas características?
- Como escalonar as ações de cobrança para ter a melhor relação custo benefício?



# Mineração de Dados Possibilita Responder às Seguintes Perguntas:

- Como planejar a expansão dos meus canais de telefonia em função do fluxo de utilização?
- Como escalonar determinados recursos de uma companhia aérea em função da demanda?
- Qual é a demanda esperada de utilização dos recursos de um hospital para pacientes que têm idade acima de 50 anos, moram em X e possui hábitos A, B e C?

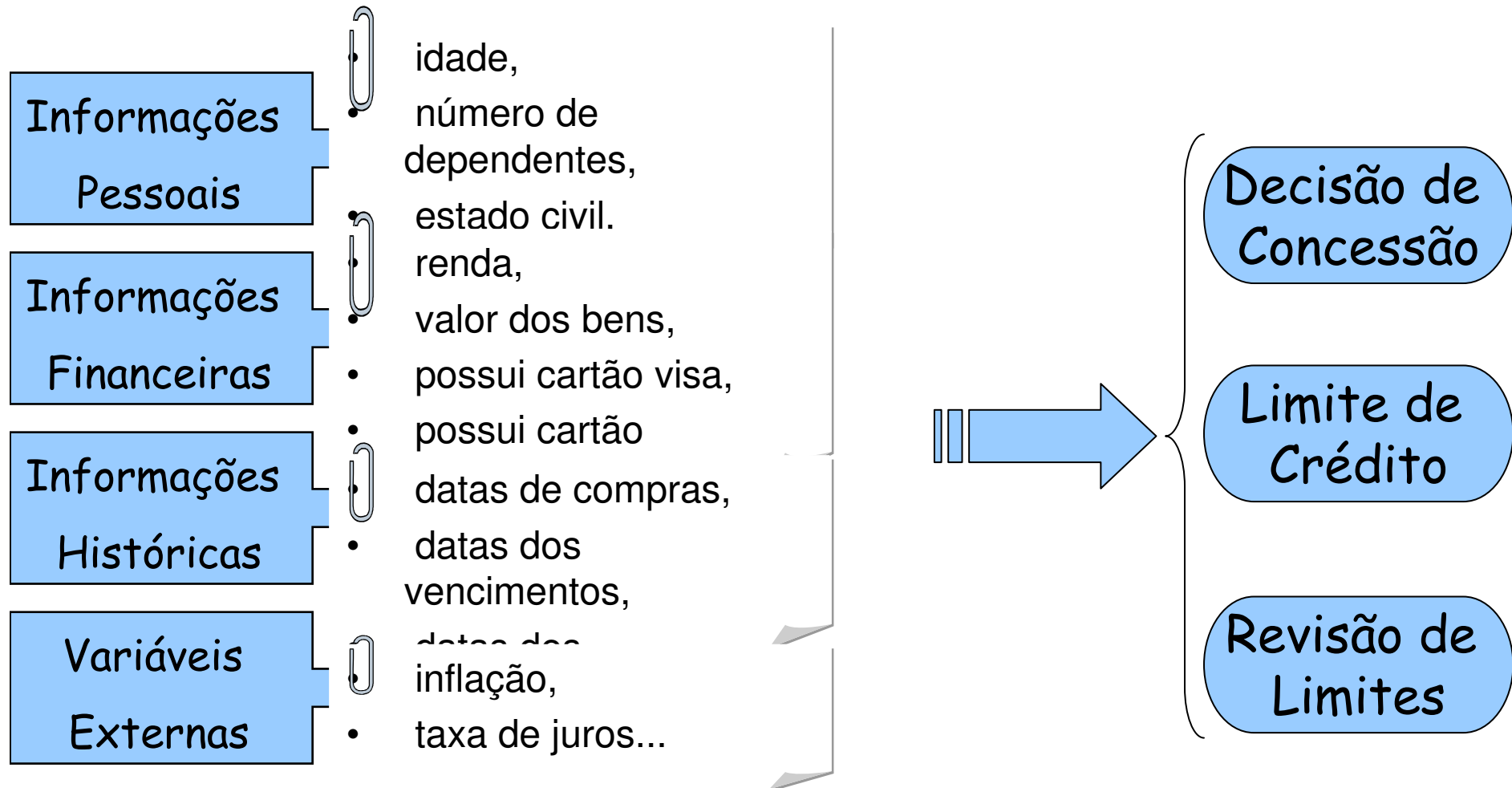
# Técnicas da Inteligência Computacional Utilizadas em Mineração de Dados

- Estatística
- Regras de associação e classificação
- *Clustering*
- Redes neurais
- Algoritmos genéticos
- Lógica difusa/nebulosa
- Inteligência artificial simbólica

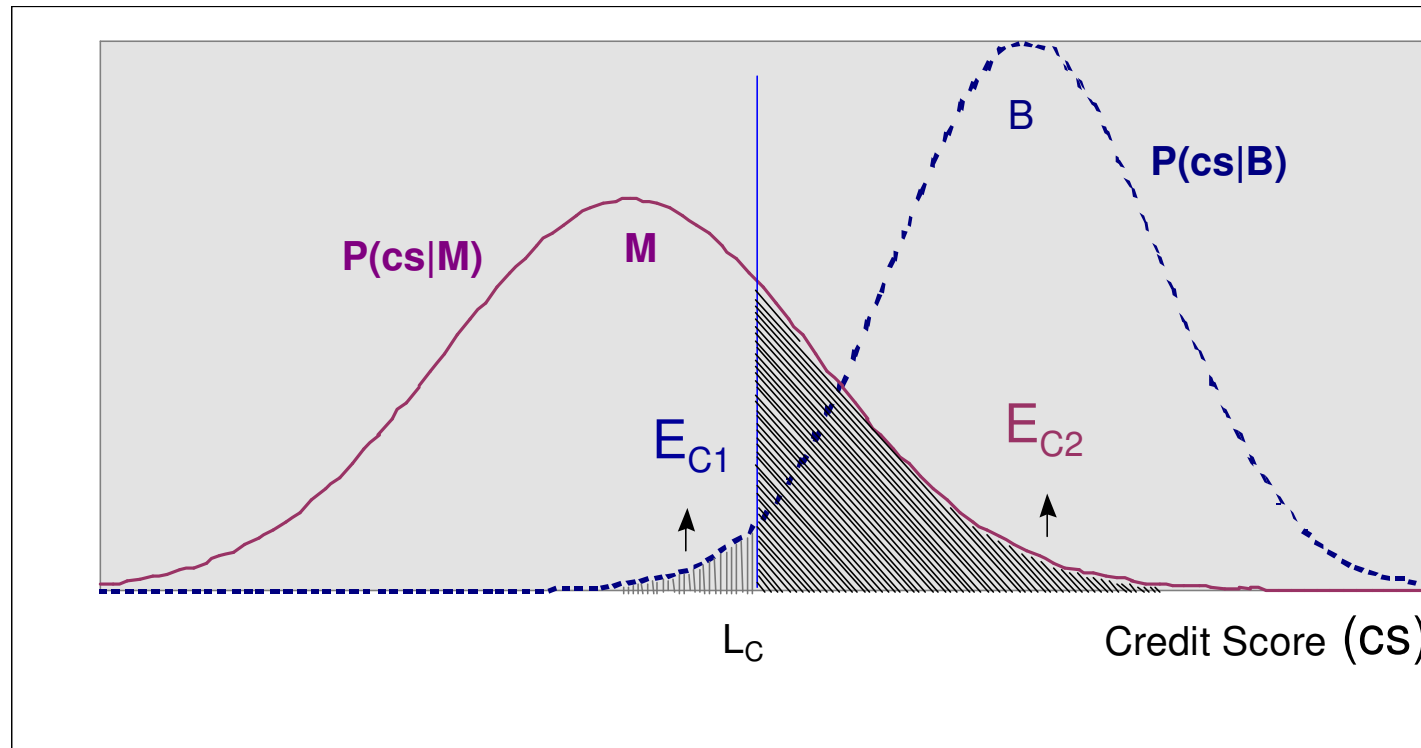
# Exemplos de Serviços na Área de Crédito

- Análise de comportamento do consumidor
  - inadimplência
  - valor do cliente (shareholders' value)
  - abandono do serviço (attrition)
- Detecção de Fraudes
  - na concessão
  - nas transações
- Cobrança
  - priorização na cobrança das dívidas
  - identificação das dívidas “podres”
- Mala direta dirigida / Identificação de nichos de mercado
  - identificação de nichos com características interessantes
  - produtos a oferecer por nicho de mercado
  - formação de kits para venda casada

# O Problema da Análise de Crédito



# Erros da Decisão



- $E_{C1}$  - Erro de classificar bons clientes como maus pagadores
- $E_{C2}$  - Erro de classificar maus clientes como bons pagadores

# Minimização do Risco de Crédito

- *Definição do limiar de decisão do score a partir da massa conhecida. Ponto de equilíbrio:*

$$L_C = cs | \sum \text{Prejuízo dos maus aceitos} = \sum \text{Lucro dos bons rejeitados}$$

- *Decisão sobre a concessão de crédito a um novo solicitante:*

$$cs \geq L_C ? \left\{ \begin{array}{l} \text{Sim} \Rightarrow \text{Concede} \\ \text{Não} \Rightarrow \text{Rejeita} \end{array} \right\} \text{ o crédito}$$

# Política Real de Risco de Crédito

- *Definição de 2 limiares de decisão do score ( $L_I$  e  $L_S$ ) a partir da massa conhecida considerando o retorno financeiro.*
- *Decisão sobre a concessão de crédito a um novo solicitante:*

$$cs \leq L_I ?$$

Rejeita o crédito

$$L_I < cs < L_S ?$$

Decisão do analista de crédito

$$cs \geq L_S ?$$

Concede o crédito

# Caso Real: Financeira Abrangência no Brasil

- Massa de Dados
  - Em torno de 500.000 registros disponibilizados
  - 250 mil disponíveis para uso e 250 mil disponíveis para validação pela Financeira
- Utilizados
  - Em torno de 60.000 registros
  - 45.000 para treinamento
  - 15.000 para teste
- Utilizados para validação pela Financeira
  - Em torno de 100.000 registros



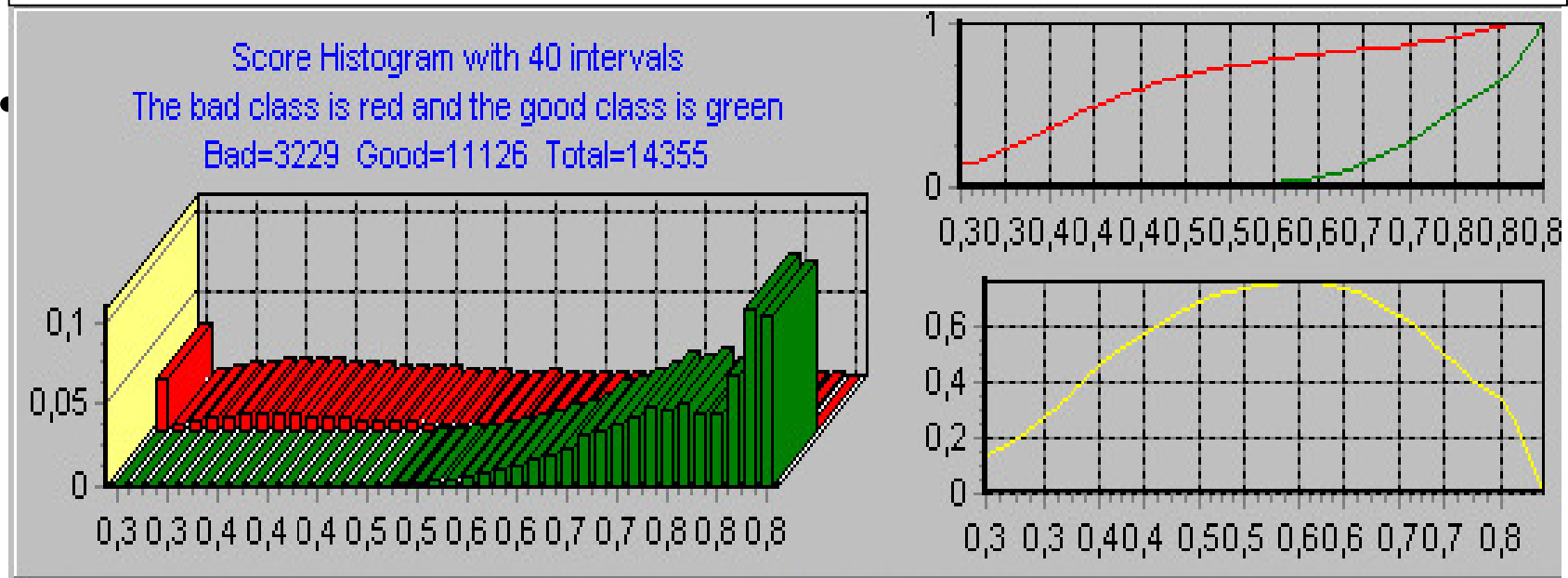
# Teste Estatístico KS (Kolmogorov-Smirnov)

- Permite a observação do nível de separação entre as distribuições de bons e maus pagadores
- Índice não-linear (entre 0 e 100)
- Quanto maior o valor de KS melhor a separação entre as distribuições
- Valores do KS no conjunto de validação (Financeira)
  - Solução do NeuralScorer : KS=39
  - Modelo atual da Financeira: KS=32

# Gráfico do KS – Conjunto de Desenvolvimento

- Permite a observação do nível de separação entre as distribuições de bons e maus

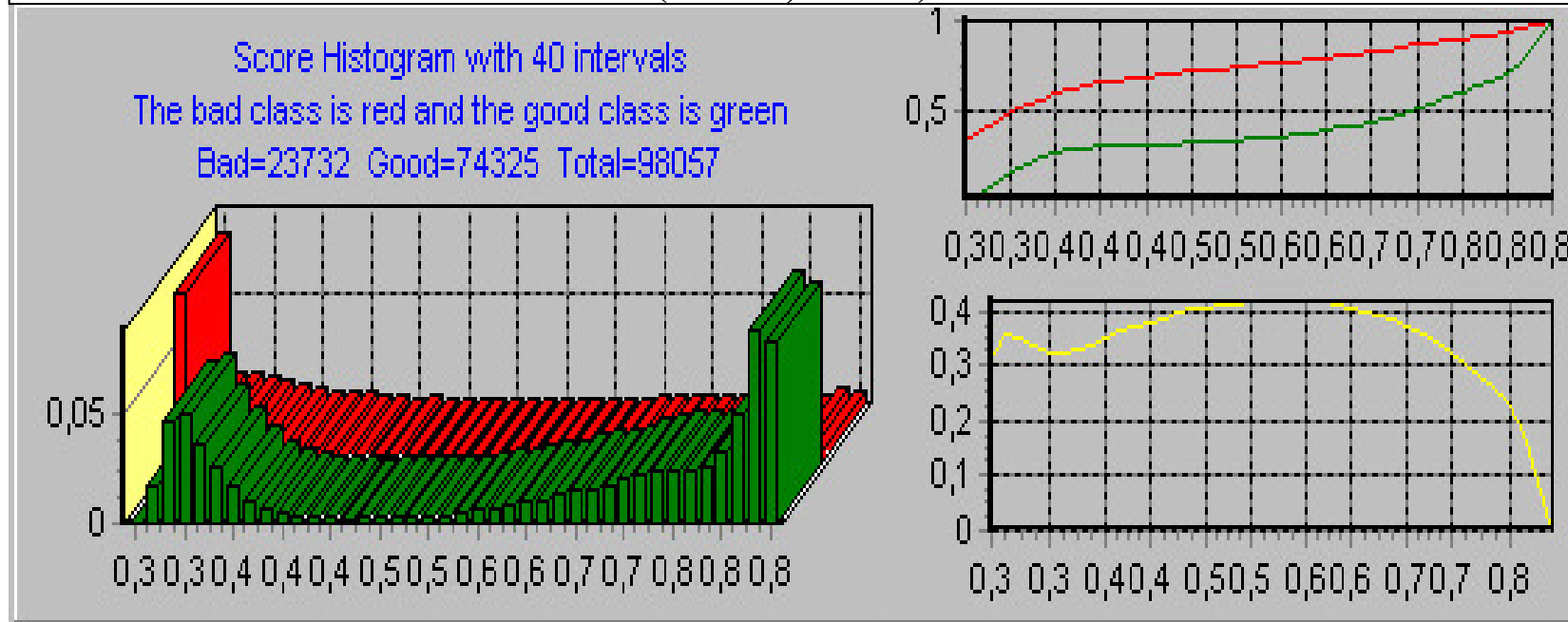
Conjunto para controle da estimação paramétrica (14355 registros)



# Gráfico do KS – Conjunto de Validação da Financeira

- Permite a observação do nível de separação entre as distribuições de bons e maus

Conjunto para controle de qualidade retido pela Financeira (98057 registros)  
(KS2=0,419448)



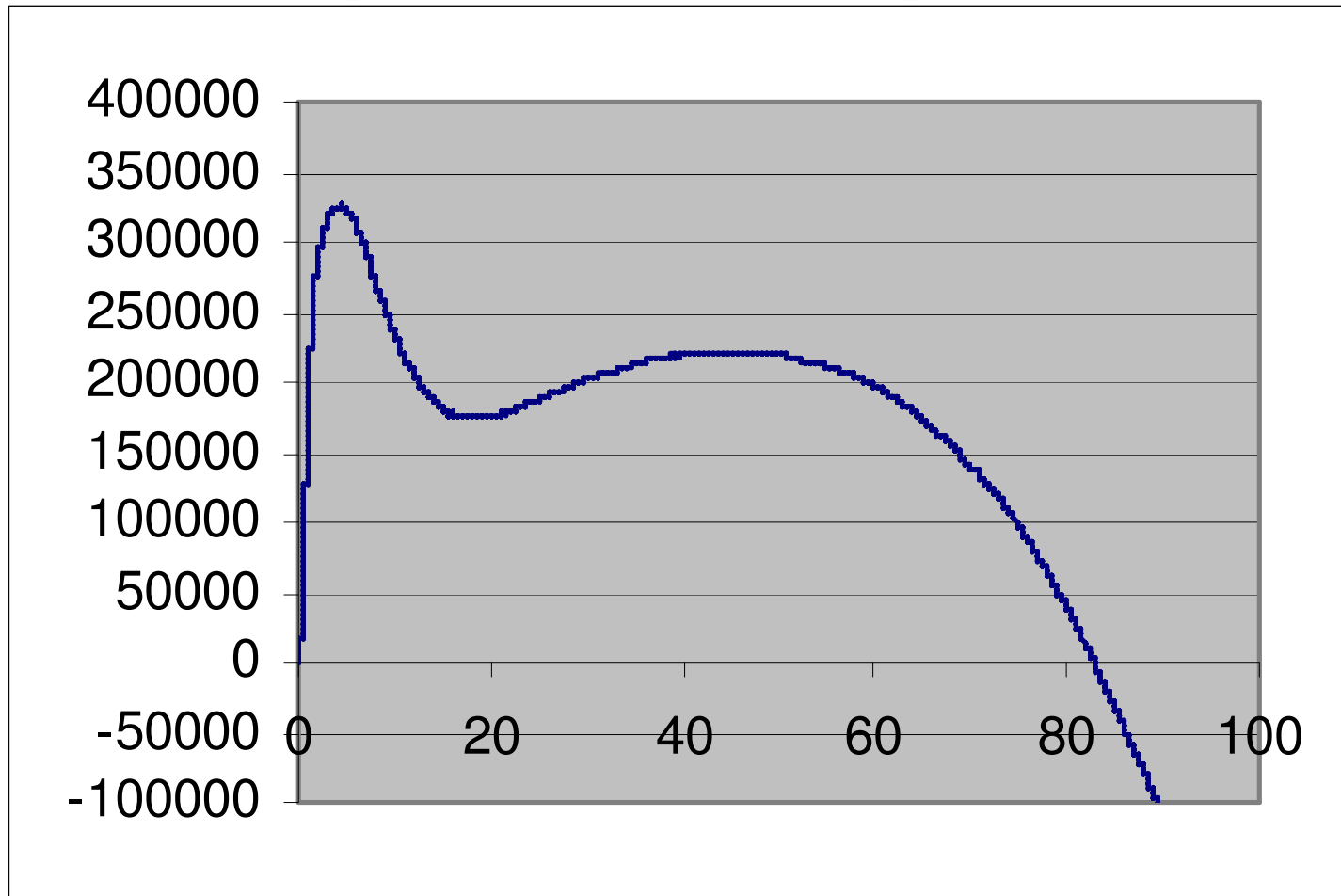
# Retorno Financeiro em Função do Score

- *Definição do Retorno Financeiro (RF) sobre uma massa conhecida:*

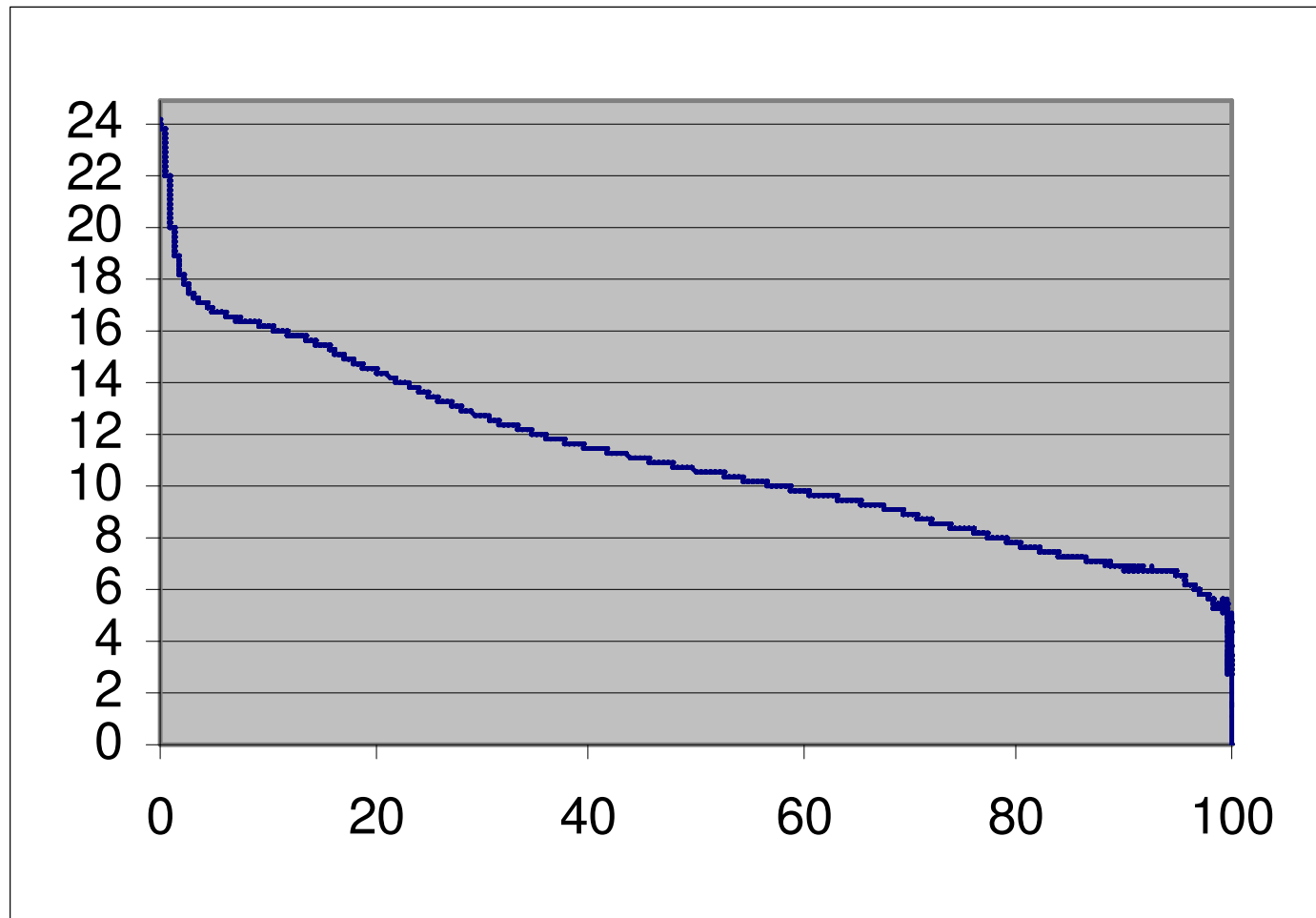
$$RF = \sum \text{Prejuízo dos maus detectados} - \sum \text{Lucro dos bons rejeitados}$$

- *Relação de custo entre o prejuízo causado pela aceitação de um mau pagador e a rejeição de um bom pagador deve ser considerada...*

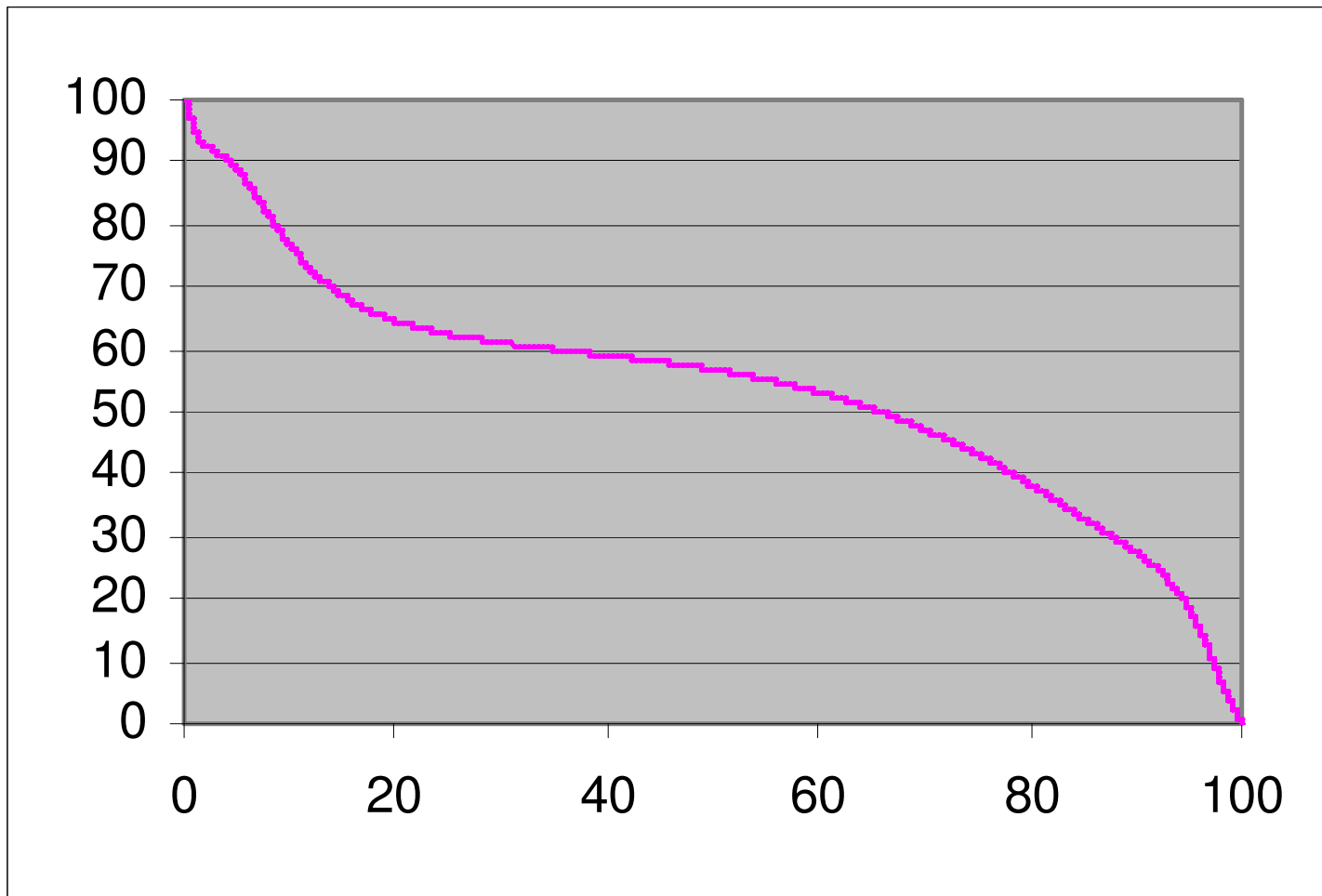
# Retorno Financeiro (relação 1/2) em Função do Score (acumulado)



# Risco de Inadimplência em Função do Score



# Massa Mantida em Função do Score



# Redução do Risco de Inadimplência

- *Situação Inicial*
  - Taxa de inadimplência: 24%
- *Solução do NeuralScorer (sugestão)*
  - Ponto de corte no score: 17
  - Taxa de inadimplência: 16%
  - Redução da taxa de inadimplência: 33%
  - Massa de bons mantida: 85%
  - Relação bons/maus: 1,88
  - Retorno financeiro (100.000 clientes)
  - R\$ 325.000,00 (relação 1/2)



# Redução do Risco de Inadimplência

- *Situação Inicial*
  - Taxa de inadimplência: 24%
- *Solução do NeuralScorer (sugestão)*
  - Ponto de corte no score: 5
  - Taxa de inadimplência: 17%
  - Redução da taxa de inadimplência: 29%
  - Massa de bons mantida: 99%
  - Relação bons/maus: 0,14
  - Retorno financeiro (100.000 clientes)
  - R\$ 230.000,00 (relação 1/2)

# **Detecção de Fraudes**

# Regras dos Fraudadores

- *SE COD\_PROFISSAO=92400 ENTÃO  
INAD=7.35% (Repres=3.24%)*
- *SE CD\_UF\_RES=GO ENTÃO INAD=7%  
(Repres=4.19%)*
- *SE CD\_SEX=2 e CD\_OCU=2 ENTÃO  
INAD=3.38% (Repres=4.82%)*

# Retorno Financeiro

## ***FraudDetector***

- *Prejuízo associado com os fraudadores para a massa de desenvolvimento*
  - R\$ = 2.036.453 (60 mil registros)
  - R\$ = 9.892.342 (280 mil registros)
    - *Desenvolvemos uma solução com o FraudDetector que reduz o prejuízo em 60%*
- R\$ = 814.581 (60 mil registros)
- R\$ = 3.956.936 (280 mil registros)

# Maiores Informações

- Germano Crispim Vasconcelos
- E-mail: [gcv@cin.ufpe.br](mailto:gcv@cin.ufpe.br)
- Homepage: [www.cin.ufpe.br/~gcv](http://www.cin.ufpe.br/~gcv)