

Estratégias de Rejeição para Problemas Multiclasses

Andréia Marini, Alessandro L. Koerich

Pontifícia Universidade Católica do Paraná, Curitiba, PR, Brasil
{andreia, alekoe}@ppgia.pucpr.br

Resumo Estratégias de rejeição são utilizadas para melhorar o desempenho de sistemas de reconhecimento de padrões. Entretanto, as opções descritas na literatura referem-se a problemas bem definidos em conjuntos de dados e objetivos específicos, em geral considerando apenas duas classes. Realizamos um estudo comparativo para caracterizar o desempenho de algumas estratégias de rejeição em problemas com múltiplas classes. Os resultados apontam que as características dos conjuntos de dados utilizados podem influenciar o comportamento dos mecanismos de rejeição e que métodos clássicos da literatura podem apresentar resultados similares a estratégias heurísticas.

Key words: Estratégias de rejeição, Problemas com múltiplas classes, Curvas ROC

1 Introdução

Nos últimos anos, várias técnicas têm sido desenvolvidas e testadas visando encontrar melhores resultados para sistemas inteligentes. A busca é direcionada em atribuir a máquina, capacidade de aprendizagem e generalização similar a de um ser humano, sendo a classificação uma das principais tarefas que fazem parte de tais técnicas. Visando incrementar o processo de classificação, o conceito de rejeição admite que um sistema de reconhecimento aplique uma decisão global de aceitar ou recusar uma hipótese se o classificador não estiver certo o suficiente. Uma evidência sobre a certeza é dada pelo valor da probabilidade atribuída às hipóteses fornecidas pelo classificador. A recusa de uma hipótese pode acontecer quando um padrão é ambíguo não podendo ser associado a uma única classe com grande certeza, ou quando este padrão pode ser associado com baixa confiança a diversas classes.

Uma estratégia de rejeição é uma ferramenta útil para melhorar a confiabilidade no processo de classificação. Em algumas aplicações, o custo de rejeitar certos padrões e controlar os procedimentos necessários após sua implementação, realizando uma classificação manual pode ser menor que o custo de uma classificação incorreta. Mesmo trabalhando com taxas de rejeição elevadas essa opção pode se tornar viável em tarefas que são realizadas, tradicionalmente, de forma

manual. Dentre as possíveis aplicações para este contexto destacamos a manipulação de vários tipos de documentos, como: formulários, recibos, cheques bancários e envelopes postais.

O interesse por estratégias de rejeição é encontrado em muitos trabalhos reportados na literatura. Uma consideração importante em relação a esses trabalhos é fato de que eles fazem referência a um único problema e procuram encontrar uma solução ótima para tal problema. A maioria deles é voltado ao comportamento de estratégias de rejeição em bases de dados de palavras ou caracteres, *on-line* [3],[8] ou *of-line* [5],[7], considerando na maioria das vezes métodos estatísticos e não considerando que parâmetros ou características do problema podem impactar nos resultados apresentados pelas estratégias de rejeição.

Algumas dessas estratégias utilizam o formalismo estatístico, como é o caso das estratégias propostas por [1] e [2]. Outras propostas interessantes são apresentadas por [9] e [4]. O primeiro sugere a construção de um sistema de reconhecimento de padrões, com múltiplos estágios e a integração de mecanismo de rejeição. O segundo apresenta um método buscando encontrar um compromisso de erro-rejeição satisfatório. Uma estratégia de rejeição medindo a confiança na saída de uma rede neural de reconhecimento de cheques bancários, através de um *score*, sendo capaz aceitar ou rejeitar os exemplos e atribuí-los a uma lista de bons e maus candidatos.

Neste trabalho propomos verificar o comportamento de diferentes estratégias de rejeição, algumas probabilísticas e outras heurísticas, considerando problemas de classificação que possuem características distintas em conjuntos de dados diferentes.

Este artigo está organizado em cinco seções. Após uma breve introdução, a Seção 2 apresenta métodos de rejeição encontrados na literatura. A Seção 3 descreve as estratégias de rejeição propostas. A Seção 4 apresenta os resultados obtidos através dos experimentos realizados. As conclusões são apresentadas na última seção.

2 Estratégias de Rejeição

Para formalizar a rejeição usamos a notação de confiança associada a uma hipótese. Consideramos um problema de classificação atribuído a um classificador C que fornece na saída uma medida de confiança c_i para cada uma das “ c ” classes pertencentes ao problema onde, $\sum_{i=1}^c c_i = 1$. Quando um exemplo é apresentado ao classificador, este deve atribuir uma medida de confiança a cada uma das c classes as quais este exemplo possa pertencer. No caso do classificador utilizado ser uma RNA, este valor refere-se a probabilidade atribuída pelo classificador sempre representado por valores reais entre 0 e 1. Sendo que quanto mais próximo de 1 maior é o grau de confiança do exemplo pertencer a classe indicada. Idealmente espera-se que para uma dada instância i teremos, $c_j = 1$ e $\sum_{i=1}^{c-1} c_i = 0$ para todo o $i \neq j$.

Posteriormente, conhecidas as medidas de confiança $(c_1, c_2 \dots c_c)$ fornecidas pelo classificador, uma estratégia de rejeição pode ser aplicada. Com base nestas medidas, uma determinada estratégia de rejeição pode decidir se concorda com o classificador sobre a classificação de um determinado exemplo, ou se a rejeita e deixa tal exemplo para um procedimento mais sofisticado, uma classificação manual ou a utilização de outro classificador mais especializado, por exemplo. A rejeição ocorre quando um padrão ambíguo, propenso a ser incorretamente classificado, é deixado de lado para uma posterior classificação. A decisão de aceitar ou rejeitar um exemplo é controlada por um limiar t . Quando um exemplo possui medidas de confiança $\hat{c} \geq t$ onde $\hat{c} = \max c_j, 1 \leq j \leq c$ ele é aceito, enquanto exemplos com $\hat{c}_j < t$ são rejeitados. O objetivo de aplicar uma estratégia de rejeição em um dado problema é encontrar um valor ótimo para T . Quando este valor é encontrado uma estratégia de rejeição consegue rejeitar todos os exemplos incorretamente classificados pelo classificador e também aceitar todos os exemplos corretamente classificados. A seguir apresentamos três estratégias de rejeição descritas na literatura:

- **Rejeição com um limiar (Chow)**- Este mecanismo é proporcionado pela regra de Chow [1]. Apenas um limiar T é utilizado para todas classes do problema. De acordo com a regra de Chow uma instância é rejeitada caso obedeça a Equação 1, por outro lado, uma instância x é aceita e associada classe caso obedeça a Equação 2.

$$\max_{k=1, \dots, N} P(\omega_k | x) = P(\omega_i | x) < T \quad (1)$$

Onde $T \in [0,1]$.

$$\max_{k=1, \dots, N} P(\omega_k | x) = P(\omega_i | x) \geq T \quad (2)$$

- **Rejeição com múltiplos limiares (Fumera)**- A reprodução desta estratégia é baseada em [2]. Múltiplos limiares são utilizados um para cada classe do problema. Para esta estratégia uma instância x é rejeitada caso obedeça a Equação 3. Enquanto, uma instância x é aceita e associada à classe ω_i , caso obedeça a Equação 4. Portanto, utilizando o mecanismo CRT para uma tarefa de classificação com N classes de dados caracterizadas por probabilidades *a posteriori* estimadas $\hat{P}(\omega_i | x)$, $i = 1, \dots, N$.

$$\max_{k=1, \dots, N} \hat{P}(\omega_k | x) = \hat{P}(\omega_i | x) < T_i \quad (3)$$

$$\max_{k=1, \dots, N} \hat{P}(\omega_k | x) = \hat{P}(\omega_i | x) \geq T_i \quad (4)$$

- **Rejeição utilizando a diferença entre os valores de confiança (DIF)**- Esta estratégia é baseada em [3]. Sua implementação é obtida calculando a diferença entre os dois maiores níveis de confiança obtidos na saída da RNA.

3 Estratégias Propostas

Neste artigo propomos três diferentes estratégias de rejeição. Dois novos métodos heurísticos (Médias das Classes e Média/Desvio) e um estatístico (FumeraMod) implementados afim de serem comparados com demais métodos já citados.

- **Rejeição com múltiplos limiares (FumeraMod)**- Esta estratégia é uma alteração a regra CRT, é baseada na idéia da estratégia gulosa (*Greedy*), ou seja, encontramos uma solução ótima para cada subproblema (Classe) pretendendo chegar a uma solução ótima global para o problema da rejeição. Neste caso, buscamos os limiares ótimos locais considerando uma classe de cada vez, para obter os limiares é necessário dividir o conjunto de validação em vários subconjuntos, sendo um para cada classe. A principal diferença em relação ao CRT é a escolha do limiar local e não global. A aplicação do mecanismo de rejeição da mesma forma que as estratégias anteriores.
- **Rejeição utilizando Média das Classes (*Class Average*)**- Esta estratégia é um método heurístico para calcular um suposto limiar ótimo a ser adotado como critério de rejeição. Aqui, o limiar é obtido de forma automática em relação ao conjunto de dados através da fórmula matemática para encontrar a média de um conjunto de valores. Neste caso, extraímos a média de um subconjunto (classe selecionada) considerando somente os exemplos corretamente classificados de uma determinada classe.
- **Rejeição utilizando Média das Classes e Desvio Padrão (*Class Average and Standard Deviation*)**- Esta estratégia é derivada da anterior. Neste caso além de calcular a média calculamos também o desvio padrão para os dados corretamente classificados. O limiar é obtido através da diferença desses dois valores.

4 Experimentos

A base de dados utilizada para o reconhecimento de caracteres foi a NIST (*National Institute of Standards and Technology*) *NCharacter - SD19* contendo imagens no formato TIFF das 52 letras do alfabeto sendo cada letra uma classe. Foram considerados três conjuntos de dados:

- **Conjunto Upper** - Utilizamos 26 classes diferentes (“A - Z”) de caracteres maiúsculos totalizando 61.473 exemplos, divididos em três conjuntos: treinamento (37.440), validação (12.092) e teste (11.941). Os resultados do classificador MLP sem aplicação de mecanismos de rejeição, ou seja, taxa de rejeição 0% apresentam taxa de reconhecimento de 97,87% para o conjunto treinamento, 93,60% para o conjunto de validação e 92,49 % para o conjunto de teste.
- **Conjunto Lower** - Utilizamos 26 classes diferentes de caracteres minúsculos (“a - z”) com 61.018 exemplos, divididos em treinamento (37.440), validação (11.578) e teste (12.000). Os resultados do classificador apresentam taxa de reconhecimento de 95,82% para o conjunto treinamento, 90,05% para o conjunto de validação e 86,73% para o conjunto de teste.

- **Conjunto UpperLower** - A combinação das bases anteriores em 52 classes, caracteres maiúsculos e minúsculos, onde “A” e “a” pertencem a mesma base e são classes distintas. Os resultados do classificador MLP apresentam taxa de reconhecimento de 80,93% para o conjunto treinamento, 73,59% para o conjunto de validação e 69,08 % para o conjunto de teste.

O classificador escolhido é uma rede neural do tipo MLP utilizada em [6]. A seguinte arquitetura foi definida para esse problema: uma única camada oculta com 100 neurônios, 108 unidades de entrada (características extraídas das imagens) e 26 unidades de saída (quantidade de classes) sendo esta arquitetura adotada para os conjuntos Upper e Lower. A arquitetura adotada para o Conjunto UpperLower é composta por uma única camada oculta com 150 neurônios, com 108 unidade de entrada (características extraídas da imagens) e 52 unidades de saída (quantidade de classes).

O principal objetivo deste trabalho é avaliar diferentes estratégias de rejeição através de curvas ROC (*Receiver Operating Characteristics*) em diferentes problemas de classificação. Para avaliar a eficiência das estratégias de rejeição usamos a terminologia convencional utilizada para reconhecimento de padrões que refere-se à falsa aceitação e falsa rejeição, tratando o correto e o incorreto reconhecimento, verificados através da implementação de rotinas de pós-processamento. A saída do pós-processamento pode ser definida de quatro formas:

- **CA: Correta Aceitação** - Instâncias corretamente reconhecida pelo classificador e após aplicação de uma estratégia de rejeição ela é novamente aceita como correta.
- **CR: Correta Rejeição** - Instâncias incorretamente reconhecida pelo classificador, e rejeitada pela estratégia de rejeição.
- **FA: Falsa Aceitação** - O classificador classifica incorretamente uma instância, mas a estratégia de rejeição aceita.
- **FR: Falsa Rejeição** - O classificador classifica corretamente uma instância, mas a estratégia de rejeição rejeita esta instância.

Tais estatísticas, são utilizadas para avaliação do desempenho de uma tarefa de pós-processamento ou simplesmente, das estratégias de rejeição. Podemos medir a precisão de tais rotinas observando [3]:

- **Taxa de Falsa Aceitação** - FAR, ou *Erro do tipo I*: Aceitar que um exemplo foi classificado corretamente quando na realidade foi classificado incorretamente. A frequência de ocorrências deste tipo é chamada de FAR calculada pela Equação 5:

$$FAR = \frac{FA}{FA + CR} \quad (5)$$

- **Taxa de Falsa Aceitação** - FRR, ou *Erro do tipo II*: Aceitar que um exemplo foi classificado incorretamente quando na realidade foi classificado

corretamente. A frequência de ocorrências deste tipo é chamada de FRR calculada pela Equação 6:

$$FRR = \frac{FR}{FR + CA} \quad (6)$$

4.1 Base NIST Caracteres Maiúsculos - (*Upper*)

A comparação do desempenho das seis estratégias de rejeição utilizadas para o conjunto Upper em termos de taxa de falsa rejeição (FRR) e taxa de falsa aceitação (FAR) pode ser vista na Figura 1. A curva ROC em cada estratégia refere-se aos exemplos das 26 classes pertencentes ao conjunto de teste, ou seja 11941 exemplos. A análise visual do gráfico ROC aponta a estratégia mais adequada para este conjunto foi a Média das Classes, provendo uma correta aceitação de 10116 dos exemplos e 686 são corretamente rejeitados. As estratégias Chow, Fumera, Média/Desvio apresentam resultados similares. A verificação das taxas de falsa aceitação e falsa rejeição apresenta para a estratégia Média das Classes valores 11,5% para FAR e 17,4% para FRR, e uma taxa de erro 0,86%. Utilizando a Média/Desvio como estratégia os valores são alterados para FAR 23,5% e FRR 8,4% para uma taxa de erro de 1,76%. As estratégias FumeraMod e DIF apresentam índices elevados de FAR em média 43,6% e índices baixos de FRR em média 2,1%, caracterizando o pior desempenho neste conjunto. Os valores referentes FAR para a estratégia Chow variam entre 0,3% e 44,8% e FRR entre 65,3% e 3,5%. Para Fumera, FAR varia de 0,2% a 44,9% e FRR entre 79% e 3,5% ambos relacionados a taxa de erro entre 0% e 5%. Resumindo, para este conjunto de dados a estratégia mais adequada é Média das Classes.

4.2 Base NIST Caracteres Minúsculos - (*Lower*)

A análise visual do gráfico ROC aponta o desempenho das seis estratégias de rejeição na Figura 2, apresentando a estratégia DIF como a mais próxima do ponto (0,0) com FAR variando de 0% até 50,5%. Verificando taxas de falsa aceitação e falsa rejeição observamos que a estratégia Média/Desvio apresenta valores de 44,2% para FAR e 10,7% para FRR, e uma taxa de erro 5,85%. Utilizando a Média das Classes como estratégia os valores são alterados para FAR 23,9% e FRR 20,7% para uma taxa de erro de 3,16%. Os valores referentes as estratégias Chow e Fumera novamente são similares, ambas apresentam FAR variando entre 1,9% e 42,3% e FRR entre 10,7% e 65,1%. A estratégia FumeraMod apresenta variação de 6,3% e 47% para FAR e 42,6% e 9,8% para FRR.

4.3 Base NIST Caracteres Maiúsculos e Minúsculos - (*UpperLower*)

A comparação do desempenho das seis estratégias de rejeição em termos de taxa de falsa rejeição (FRR) e taxa de falsa aceitação (FAR) para o conjunto UpperLower pode ser vista na Figura 3. A análise visual do gráfico ROC aponta a estratégia DIF com melhor desempenho e a estratégia Média/Desvio com o

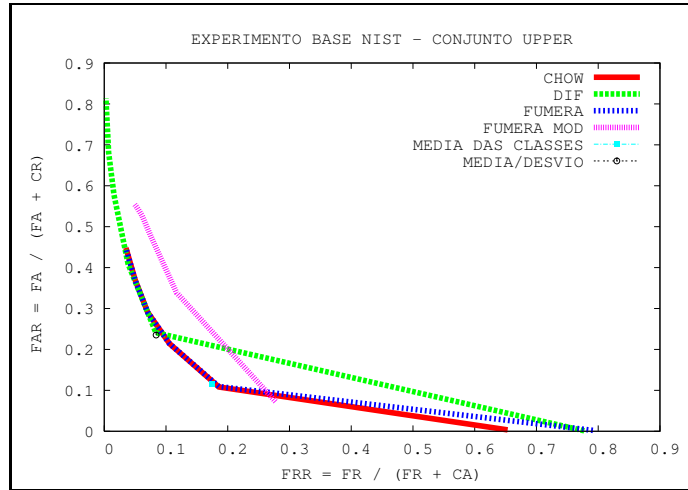


Figura 1. Avaliação das estratégias de rejeição utilizando Curva ROC para o conjunto Upper.

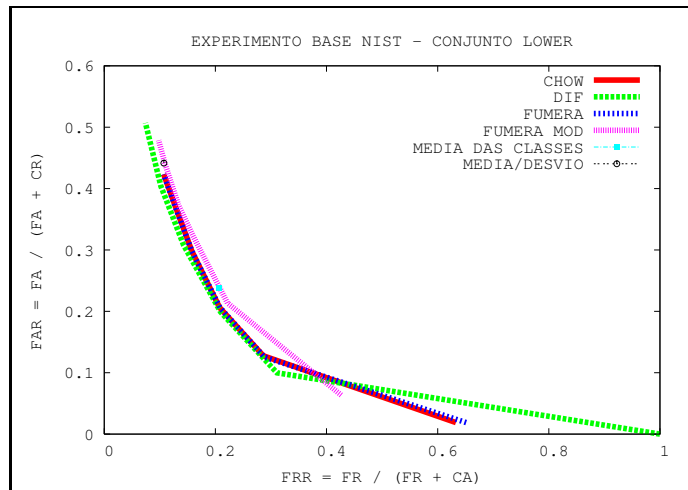


Figura 2. Avaliação das estratégias de rejeição utilizando Curva ROC para Base Lower.

pior desempenho, concordando com os resultados encontrados na avaliação do conjunto anterior. Depois da estratégias DIF os melhores desempenhos são apontados por Chow e Fumera seguidos de FumeraMod e Média das Classes apresentando resultados similares. A verificação das taxas de falsa aceitação e falsa rejeição apontam a estratégia Média/Desvio apresentando 55,82% para FAR e 27,57% para FRR, e uma taxa de erro 16,60%, ou seja o pior desempenho para o conjunto. Utilizando a Média das Classes como estratégia os valores são alterados para FAR e 36,2% e FRR 30,9% para uma taxa de erro de 30,91%. A estratégia FumeraMod apresenta FAR entre 1,6% e 45,5% e FRR entre 79,2 % e 21,3%. Os valores referentes FAR para a estratégia Chow variam entre 0,8% e 22,3% e FRR entre 82,7% e 36,2%. Para Fumera, FAR varia de 0,6% a 22,3% e FRR entre 84,6% e 32,2% ambos variando a taxa erro entre 0% e 5%.

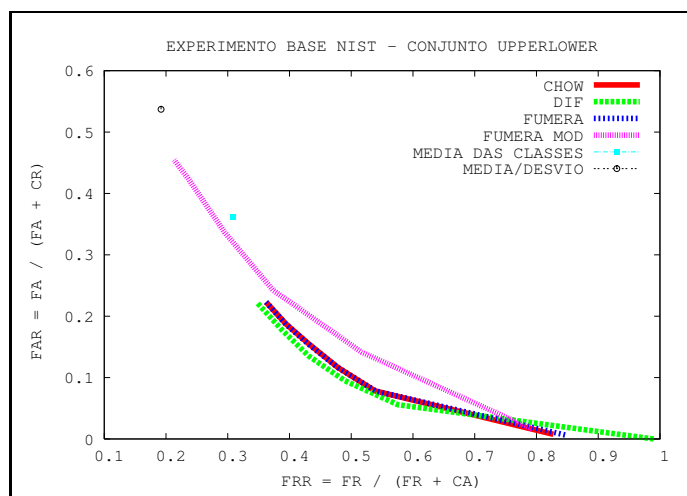


Figura 3. Avaliação das estratégias de rejeição utilizando Curva ROC para Base UpperLower.

Tabela 1. Resumo do comportamento das estratégias de rejeição considerando a análise ROC para os três conjuntos de dados utilizados.

Conjunto Upper	Conjunto UpperLower	Conjunto Upper
Média das Classes	DIF	DIF
Chow	Chow ou Fumera	Chow ou Fumera
Fumera	FumeraMod	FumeraMod
Média/Desvio	Média das Classes	Média das Classes
DIF	Média/Desvio	Média/Desvio
FumeraMod		

4.4 Análise dos resultados

A regra heurística representada pela estratégia DIF obteve desempenho na maioria dos casos melhor ou igual a estratégias estatísticas ditas como "ótimas" na literatura, sendo o caso do Chow e Fumera. A idéia de encontrar os limiares ótimos locais em cada classe representada pela estratégia FumeraMod apresentou na maioria dos casos como uma alternativa inadequada, sugerindo que nem sempre um ótimo local é também um ótimo global, pois, outros métodos com menores custos de implementação apresentam resultados mais adequados. As demais regras heurísticas representadas pelas estratégias Média das Classes e Média/Desvio também apresentam resultados viáveis para aplicação em um sistema de reconhecimento, principalmente pelo fato do processo de obtenção dos limiares para cada classe apresenta-se de forma simples e automática. Entretanto, parecem ser as estratégias que mais são influenciadas pela formação do conjunto de dados. A Tabela 1 apresenta as estratégias ordenadas de acordo com a distância em relação ao ponto (0,0) dos seus respectivos gráficos em conjuntos de dados com múltiplas classes. Analisando a tabela percebemos que a análise ROC aponta a mesma ordem de desempenho para as estratégias implementadas para os conjuntos Lower e UpperLower. Apontando a seguinte sequência: DIF com o melhor desempenho, Chow e Fumera com resultados equivalentes, seguidos de FumeraMod, Média das Classes e Média/Desvio. Nestes conjuntos a dificuldade do classificador evidente nas taxas sugeridas pela RNA-MLP. Percebemos que a estratégia DIF destacou-se principalmente trabalhando com problemas de classificação mais complexos. Quando esta estratégia foi utilizada por [3] seu desempenho em relação as demais não se destacou. Neste caso, eles aproveitam a flexibilidade de uma tarefa de pós-processamento e realizaram teste com oito diferentes estratégias relacionadas as medidas de confiança oferecidas pelo classificador. Acredita-se que o pior desempenho atribuído à estratégia Média/Desvio deve-se ao fato desta estratégia ser diretamente indexada pelo desvio padrão influenciado pelas características dos conjuntos Lower e UpperLower. De forma geral, todas as estratégias melhoraram significativamente o desempenho do sistema de reconhecimento, pois, as taxas de erros diminuíram e as taxas de acerto aumentaram.

5 Conclusão

Neste artigo realizamos uma análise de estratégias de rejeição estatísticas e heurísticas. Seis diferentes estratégias foram consideradas. O resultados mostram que estratégias clássicas na literatura como Chow e Fumera obtiveram resultados similares, em muitos casos até inferiores aos métodos heurísticos e estatísticos implementados. A estratégia DIF obteve desempenho na maioria dos casos melhor ou igual a estratégias estatísticas ditas "ótimas" na literatura. Apresentou-se como uma estratégia pouco influenciada pelas condições do conjunto de dados e ainda sua implementação não possui grau de dificuldade elevado. Enfatiza-se, que as porcentagens de erro obtidas no conjunto de teste podem ainda ser ajus-

tadas para as mesmas fixadas no conjunto de validação, estabelecendo novos critérios de obtenção dos limiares para cada estratégia.

Analisando os resultados podemos concluir que o uso das estratégias de rejeição conseguiram produzir sistemas classificadores confiáveis, além de melhorar significativamente os resultados medidos em relação a taxa de erro apresentada pelo classificador de base. Um mecanismo de rejeição ótimo pode ser aplicado para resolver diversos problemas onde se faz necessário sistemas confiáveis. Nos dias atuais esta preocupação vai desde evitar fraudes e falsificação em documentos, implantação de sistemas de identificação pessoal baseados em características biométricas, até melhorar a precisão de sistemas de produção automatizados que buscam realizar seus processos de forma mais rápida e segura.

Referências

1. C. K. Chow, On Optimum Recognition Error and Reject Tradeoff, In *IEEE Transactions in Information Thoery*, 16, 41-46, 1 (1970)
2. G. Fumera and F. Roli and G. Giacinto, Reject option with multiple thresholds, *Pattern Recognition Letters*, 33, 2099-2101, 12 (2000)
3. J. F. Pitrelli and M. P. Perrone, Confidence Modeling for Verification Post-Processing for Handwriting Recognition, In *Proc. 8th International Workshop on Frontiers in Handwriting Recognition*, 30-35, Niagara-on-the-Lake, Canada, 2002., IEEE Computer Society (2002)
4. N. Gorski, Optimizing error-reject trade off in recognition systems, *Proc. 4th International Conference Document Analysis and Recognition*, 1092-1096, Ulm, Germany, IEEE Computer Society (1997)
5. A. L. Koerich, Rejection Strategies for Handwritten Word Recognition, *IWFHR*, pp. 479-484, Ninth International Workshop on Frontiers in Handwriting Recognition (IWFHR'04) (2004)
6. A. L. Koerich, Unconstrained Handwritten Character Recognition Using Different Classification Strategies, In *Proc. of the IAPR TC3 International Workshop on Artificial Neural Networks in Pattern Recognition*, Firenze, Italy (2003)
7. M. Zimmermann and R. Bertolami and H. Bunke, Rejection Strategies for Offline Handwritten Sentence Recognition, *ICPR '04: Proceedings of the Pattern Recognition, 17th International Conference on (ICPR'04)*, 2, 550-553, Washington, DC, USA, IEEE Computer Society (2004)
8. S. Marukatat and T. Artires and P. Gallinari and B. Dorizzi, Rejection measures for handwriting sentence recognition, In *Proc. 8th International Workshop on Frontiers in Handwriting Recognition*, pages 24-29, Niagara-on-the-Lake, Canada, (2002).
9. P. Pudil and J. Novovicova and S. Baha and J. Kittler, Multistage pattern recognition with reject option, *Proc. 11th Int. Conf. Pattern Recognition*, The Hague, The Netherlands, IEEE Computer Society Press, 2, 92-95, The Hague, Netherlands, IEEE Computer Society Press (1992).