

Classificação com Opção de Rejeição Aplicado à Detecção de Erros em Canais de Comunicação

Wellington Dantas de Almeida
Departamento de Engenharia de Teleinformática
Universidade Federal do Ceará - UFC
Fortaleza, Ceará, Brasil
wellingtondantas.almeida@gmail.com

Ajalmar Rêgo da Rocha Neto
Departamento de Engenharia da Computação
Instituto Federal do Ceará - IFCE
Fortaleza, Ceará, Brasil
ajalmar@gmail.com

Resumo—A classificação de padrões é uma das principais tarefas da aprendizagem de máquina e neste artigo usamos classificadores de padrões no papel de equalizadores de canais de comunicação, com o potencial da aplicação da técnica de opção de rejeição, que busca melhorar o desempenho de um classificador binário ao impedir a classificação com alto risco de erro. Além disso, como resultado principal, aplicamos o método de opção de rejeição para automatizar o processo de detecção de erros, sem o uso das abordagens convencionais dos códigos detectores de erros, que inserem informações redundantes nos pacotes de dados. Por meio de dados sintéticos da equalização de canais e em diferentes configurações de ruído, apresentamos resultados promissores de desempenho para os classificadores com opção de rejeição aplicados na detecção de erros.

Index Terms—Classificação de padrões, Opção de rejeição, Equalização de canais, Detecção de erro.

I. INTRODUÇÃO

Um sistema de comunicação de dados deve transmitir um sinal do transmissor para o receptor por meio de um canal sem nenhuma alteração na forma do sinal, porém canais ideais dificilmente existem. No processo real de comunicação, a transmissão de informações estará inevitavelmente sujeita a um certo grau de interferência [1]. Geralmente, os canais introduzem algum tipo de alteração nos sinais transmitidos fazendo com que a informação seja erroneamente entendida na recepção. Sendo os principais fatores de alteração dos sinais a interferência intersimbólica e o ruído.

Antes do envio de um pacote de dados, é normal no processo de codificação da mensagem de dados a inserção de bits redundantes para o controle de erros. Na área de detecção de erros, algoritmos como o *Single Parity Check* (SPC) tem a função de detectar bits errados na recepção dos dados. Nesta etapa de recebimento dos pacotes de dados, os pacotes codificados são novamente verificados com detectores de erros e comparados. Se estiverem corretos, o pacote de dados continuará a ser recebido de acordo com o protocolo. Se houver um erro, o pacote deve ser enviado novamente.

O uso de classificadores de padrões na equalização de canais de comunicação já foi abordado em muitos trabalhos [2]–[4], porém não são encontrados na literatura métodos usando classificação com opção de rejeição para detecção de erros em sistemas de comunicação, onde neste método, não é necessário a inserção de bits redundantes.

A classificação com opção de rejeição foi proposta por [5], onde apresenta um classificador ideal que pode ser descrito pelo equilíbrio entre a taxa de rejeição e a taxa de acerto. Também, com a classificação com opção de rejeição, uma regra de decisão é considerada ótima quando uma determinada taxa de rejeição (ou probabilidade de rejeição) minimiza a taxa de erro (ou probabilidade de erro) dado um valor de custo de rejeição. A opção de rejeitar padrões é um processo em que se opta por não classificar (rejeitar) determinados padrões que possuem uma maior chance de serem incorretamente classificados, de acordo com algum critério definido. Os padrões rejeitados são enviados para a avaliação de um especialista humano ou são sujeitos a outros métodos. Isso aumenta a possibilidade de que padrões não rejeitados possam ser classificados corretamente.

A classificação com opção de rejeição está sendo aplicada em muitas áreas de pesquisa. Exemplos podem ser vistos para aplicações médicas [7], em classificação de imagens [8], em [9] o método aprende a região de rejeição em dados complexos, em [10] é aplicado o método para previsão de defeitos de software, em [11] é aplicado a rótulos de aplicativos onde a precisão do classificador não é satisfatória, em [12] apresenta um classificador de notas e avalia a confiabilidade da nota pela probabilidade da classe vencedora através de um valor de limite de rejeição.

O objetivo deste trabalho é avaliar o uso de classificadores de padrões com opção de rejeição para fins de detecção de erros de transmissão em canais de comunicação. Para tal, são aplicados três classificadores de padrões baseados em redes neurais: *MultiLayer Perceptron* (MLP); redes *Radial Base Function* (RBF) e *Extreme Learning Machine* (ELM). Os objetivos específicos são: (i) aplicar classificadores de padrões no problema de equalização de canais de comunicação; (ii) avaliar a aplicabilidade da técnica de classificação com opção de rejeição na detecção de erros de comunicação; (iii) avaliar o desempenho dos classificadores ao utilizar o processo de classificação com opção de rejeição, quando aplicados em problemas binários desta natureza;

O restante deste artigo está organizado da seguinte forma: na Seção II é apresentado conceitos de equalização de canais de comunicação; na Seção III é mostrado o processo de codificação dos bits e detecção de erro; na Seção IV são apre-

sentados os algoritmos de aprendizado de máquina baseados em redes neurais; na Seção V é introduzido a classificação com opção de rejeição; na Seção VI são descritas as configurações utilizadas nas simulações; na Seção VII são apresentados os resultados alcançados e na Seção VIII é apresentada a conclusão.

II. EQUALIZAÇÃO DE CANAIS DE COMUNICAÇÃO

Quando um sinal passa por um canal introduz uma distorção de amplitude e é recomendável a utilização de um equalizador. A função do equalizador é compensar a resposta não ideal de um canal de comunicação com o intuito de restaurar ao máximo o sinal transmitido [13], [14]. Na Figura 1, é apresentado um modelo de comunicação, onde para um sinal $x(n)$ enviado da transmissão passa pelo canal e tem a sua forma modificada $h(n)$, além de ser adicionado o ruído $\eta(n)$. O sinal recebido pelo equalizador, $u(n) = h(n) + \eta(n)$, é soma das duas partes. O sinal filtrado na equalização é representado por $y(n)$.

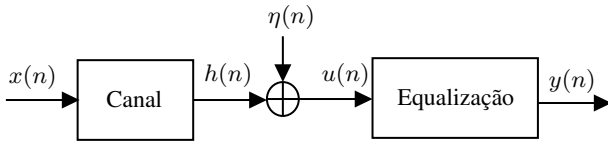


Figura 1. Modelo de um sistema de comunicação de dados que passa por um canal e é adicionado de ruído.

O equalizador pode ser representado pela filtragem adaptativa que pode ser usada para resolver vários problemas como: no sistema de cancelamento de eco telefônico, ruído, equalização de canais de comunicação, na previsão de sinal e muitos outros casos [15], [16]. Um filtro adaptativo possui características de autoajuste. Os coeficientes dos filtros adaptativos se ajustam de forma contínua e automática ao sinal dado, de modo a obter a resposta desejada a partir de algum algoritmo de minimização do erro.

Matematicamente, a equação de um filtro adaptativo pode ser representado por

$$y(n) = \sum_{i=0}^{L-1} w_i(n)x(n-i) = \mathbf{w}^T(n)\mathbf{x}(n), \quad (1)$$

onde $w_i(n)$ é o coeficiente de peso ajustável. O vetor de coeficientes $w(n)$ consiste em coeficientes de ordem L expressos como (2) e o vetor de sinal $x(n)$ no tempo n é dado por (3):

$$\mathbf{w}(n) = [w_0(n), w_1(n), \dots, w_{L-1}(n)], \quad (2)$$

$$\mathbf{x}(n) = [x(n), x(n-1), \dots, x(n-L+1)]. \quad (3)$$

III. DETECÇÃO DE ERRO

Uma condição muito comum em transmissões de dados são a ocorrência de erros, isto é, bit 0 que muda para bit 1, ou o contrário. Os sinais digitais sofrem com ruído que pode introduzir erros nos bits binários que viajam do remetente para o receptor. Para adicionar proteção contra erros em tais casos, são empregados códigos de detecção de erro, como o SPC.

A. Single Parity Check (SPC)

O código simples de verificação de paridade é um dos códigos de detecção de erros mais simples e fáceis de implementar. Resumidamente, um bit extra é adicionado a cada unidade de dados (por exemplo, byte, palavra, bloco, segmento, quadro, célula e ou pacote) do código original, isto é, k bits é alterada para uma palavra de código de n bits, em que $n = k + 1$. De forma que a nova palavra de código exibe sempre um número par ou ímpar de 1's, com o objetivo de tornar a sequência gerada (incluindo o bit de paridade) uniforme [17].

Na implementação do SPC, a paridade par (onde o número total de 1's deve par) é normalmente usada para transmissão síncrona e paridade ímpar (onde o número total de 1's deve ímpar) para transmissão assíncrona [18]. A distância mínima de Hamming para esta categoria é $D_{min} = 2$, por consequência, o código pode detectar um erro, porém não pode corrigi-lo. Para determinar se ocorreu um erro, o decodificador usa uma equação de verificação de paridade que corresponde a uma adição de módulo 2 (XOR), como mostrado em (4) e (5):

$$c_1 \oplus c_2 \oplus c_3 \oplus c_4 \oplus c_5 \oplus c_6 \oplus c_7 \oplus c_8 \oplus = 1 \quad (\text{paridade ímpar}) \quad (4)$$

$$c_1 \oplus c_2 \oplus c_3 \oplus c_4 \oplus c_5 \oplus c_6 \oplus c_7 \oplus c_8 \oplus = 0 \quad (\text{paridade par}) \quad (5)$$

onde somente será verdadeira a mensagem que houver um número ímpar de 1's em c para paridade ímpar e somente será verdadeira a mensagem que houver um número par de 0's em c para paridade par. O grande problema desta técnica de detecção de erro é que se a quantidade de erros for diferente da sua implementação, o código SPC erroneamente avalia o pacote como certo.

IV. APRENDIZADO DE MÁQUINA

A. Multilayer Perceptron (MLP)

As redes Perceptron de Multicamadas são compostas por três conjuntos de camadas: uma camada de entrada; uma ou mais camadas intermediárias (ocultas) e uma camada de saída. Na rede MLP a camada de entrada recebe os atributos do padrão. Na camada oculta os neurônios utilizam funções não-lineares e fazem interconexões entre si. Na camada de saída também é composta por neurônios que podem ser lineares (ou não), esta camada devolve o(s) valor(es) que se espera encontrar [19].

No processo de treinamento do MLP é utilizado o algoritmo de retropropagação do erro (*backpropagation*) que teve como autor principal o Rumelhart [20]. O treinamento se dá em dois sentidos: um direto e outro reverso. O direto, que consiste no cálculo das ativações dos neurônios da camada oculta que percorre a rede da camada de entrada a saída. No reverso que percorre a rede da camada de saída a camada oculta, e consiste no cálculo dos gradientes locais dos neurônios da camada de saída.

B. Radial Base Function (RBF)

A rede de base radial (RBF) é caracterizada por uma arquitetura feedforward e consiste basicamente em uma camada de entrada, uma única camada oculta e uma camada de saída [21]. Formalmente, para uma determinada entrada x , a saída de rede y pode ser escrita como [22]

$$y = \sum_{i=1}^N \omega_i R_i(x) + \omega_o, \quad (6)$$

onde ω_i são pesos, ω_o é um termo de bias, N representa o número de neurônios na camada oculta, enquanto R_i são as funções de ativação, dadas por

$$R_i(x) = \varphi(\|x - c_i\|). \quad (7)$$

Sendo φ é a função radial que fornece a característica não linear do modelo e c_i são os chamados centros RBF.

Um das vantagens da função de ativação de base radial, é possuir uma classificação robusta e velocidade de aprendizado rápida. Sua arquitetura mais comum tem somente uma camada intermediária, porém versões com mais de uma camada já foram propostas [23]. As redes RBF são aproximadores universais de funções, assim como as redes MLP, mas que apresentam um tempo de treinamento geralmente mais rápido. A rede RBF constrói aproximações locais para o mapeamento entrada e saída, enquanto o MLP constrói aproximações globais [24].

C. Extreme Learning Machine (ELM)

Esta rede foi desenvolvida para fornecer uma alternativa às redes neurais convencionais em termos de treinamento. Entre os principais detalhes do ELM, o algoritmo permite treinar uma rede neural que possui uma camada de entrada e uma camada oculta, sem o processo iterativo de retropropagação, porque os pesos são inicializados aleatoriamente na camada oculta. Desse modo, os pesos da camada de saída são estabelecidos analiticamente sem o processo de configuração de parâmetros [25]. O algoritmo oferece vantagens como velocidade de aprendizado rápida, fácil implementação e menos recursos [25].

Dado L o número de neurônios da camada oculta, N o número de padrões, β_i pesos da camada de saída e g_i a função de ativação, em (8) representa a modelagem matemática do ELM, portanto $j = 1, \dots, N$.

$$t_j = \sum_{i=1}^L \beta_i g_i(x_j) = \sum_{i=1}^L \beta_i g_i(w_i \cdot x_j + b_i) \quad (8)$$

Aqui, t_j é o vetor de saída, para um padrão de entrada x_j ; w_i e b_i são vetores de peso definidos aleatoriamente.

V. CLASSIFICAÇÃO COM OPÇÃO DE REJEIÇÃO

Classificação com opção de rejeição é uma estratégia que consiste em categorizar os padrões que possuem alta possibilidade de serem classificados incorretamente. Esses padrões

que são considerados duvidosos podem ser rejeitados e posteriormente avaliados por algum método ou por um especialista humano que pode classificá-los com mais precisão. Dessa forma, aumenta a chance dos padrões não-rejeitados serem classificados corretamente [9], [26] e [27].

Quando o conjunto de classes é constituído por somente duas saídas (-1) e ($+1$), dizemos que o problema é do tipo binário. Uma tarefa da classificação binária é definir uma saída como sendo $y = -1$ ou $y = +1$, isso com base em estimativas de probabilidades a posteriori $p(C_i|\mathbf{x})$, para um dado vetor de entrada X_i do conjunto de padrões.

Na classificação binária convencional, o espaço das amostras é dividido em duas regiões pelo classificador desenvolvido, como pertencentes a uma das classes $C1$ ou $C2$. Na classificação binária com opção de rejeição, o espaço das amostras são classificados como pertencentes a uma das classes $C1$, $C2$ ou em uma classe de rejeição, onde devem estar os objetos considerados de difícil classificação, através de algum valor de limiar de decisão. Na Figura 2 temos a representação geométrica de uma classificação binária convencional Figura 2.(a) e Figura 2.(b) com opção de rejeição em um espaço \mathbb{R}^2 . Ainda considerando a Figura 2, para padrões muito próximos do valor de decisão, há uma grande probabilidade de ocorrência de erros de classificação, note os padrões em preto, no qual o padrão pode ser atribuído incorretamente a uma classe. Com isso, a área de rejeição reserva um espaço para aumentar a confiabilidade da classificação.

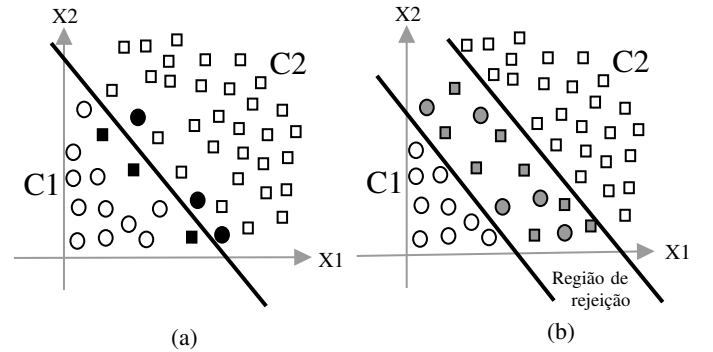


Figura 2. (a) Classificação binária convencional que divide os espaços em duas regiões: $C1$ ou $C2$; (b) classificação binária com opção de rejeição que divide os espaços em três regiões: $C1$, $C2$ ou Rejeição.

Existem três abordagens para classificação binária com opção de rejeição:

1) *Um único classificador padrão*: um padrão é rejeitado se o máximo das duas probabilidades a posteriori, ou seja, $\max\{p(\mathbf{x}|C_{-1}), p(\mathbf{x}|C_{+1})\}$, for menor que um determinado limiar. Se o classificador não fornecer saídas probabilísticas utiliza-se um limiar de rejeição. A região de rejeição é determinada após o treinamento do classificador, definindo valores limite adequados sobre a saída do classificador.

2) *Dois classificadores independentes*: nessa abordagem são utilizados dois classificadores. O primeiro classificador é treinado para classe $C1$, somente quando a probabilidade

VI. EXPERIMENTOS

de C_{-1} é elevada, o segundo classificador é treinado para saída C_2 , somente quando a probabilidade C_{+1} é elevada. A simplicidade desta estratégia é rejeitar sempre que os classificadores discordarem.

3) *Um único classificador com opção de rejeição incorporado*: essa abordagem tem a vantagem de determinar a região de rejeição durante a fase de treinamento, mas necessita que o algoritmo de aprendizagem tenha suporte a essa abordagem. Por conta disso, consiste em projetar classificadores que trazem embutido no processo de otimização de suas funções-custo a capacidade de classificação com opção de rejeição. Na abordagem com opção de rejeição embutida, o modelo aprende intrinsecamente a decidir entre as três saídas possíveis, C_{-1} , rejeição e C_{+1} .

A. Regra da Decisão Ótima

Chow possui várias contribuições relacionadas a este tema. Em [5] descreve que uma regra de decisão é ideal se dado uma taxa de erro (ou probabilidade de erro) minimiza a taxa de rejeição (ou probabilidade de rejeição). Além disto, em [6], a regra de decisão ótima deve rejeitar um determinado padrão, se o maior valor das probabilidades a posteriori for menor do que algum limiar de rejeição.

Mais explicitamente, a regra de decisão ótima δ é dada como

$$(a) \quad \delta(d_k|\mathbf{x}) = 1 \quad (k \neq 0), \quad (9)$$

em que d_k representa a decisão de aceitar o padrão \mathbf{x} e atribuí-lo à classe $k = 1, \dots, c$, tal que c representa o número de classes do problema. A Equação (9) considera que se aceite o padrão \mathbf{x} para fins de reconhecimento e identifique-o como pertencente à k -ésima classe, se

$$p_k F(\mathbf{x}|C_k) \geq p_j F(\mathbf{x}|C_j) \quad \text{para todo } j = 1, 2, \dots, c, \quad (10)$$

e

$$p_k F(\mathbf{x}|C_k) \geq (1-t) \sum_{i=1}^c p_i F(\mathbf{x}|C_i). \quad (11)$$

Em,

$$(b) \quad \delta(d_o|\mathbf{x}) = 1 \quad (k = 0), \quad (12)$$

d_o representa a decisão de rejeitar o padrão \mathbf{x} sempre que

$$\max_i \{p_i F(\mathbf{x}|C_i)\} < (1-t) \sum_{i=1}^c p_i F(\mathbf{x}|C_i). \quad (13)$$

onde \mathbf{x} é o vetor padrão, c é o número de classes, (P_1, P_2, \dots, P_c) é a distribuição de probabilidade a priori das classes, $F(\mathbf{x}|i)$ é a densidade de probabilidade condicional para \mathbf{x} , dada a i -ésima classe, $d_i (i \neq 0)$ é a decisão de que \mathbf{x} é identificado como da i -ésima classe enquanto d_o é a decisão de rejeitar, e t é uma constante entre 0 e 1 ($0 \leq t \leq 1$).

Para gerar a base de dados artificial para treinar e testar o modelo foram gerados aleatoriamente 10000 bits de dados, que formam também 10000 padrões. Uma sequência binária é formada por valores de bits 0 ou 1. Cada valor binário pode ser representado por um vetor $\mathbf{X} = \{x_1, x_2, x_3, \dots, x_m, \dots, x_i\}$ onde cada valor x_m representa uma amostra de tempo fixo do sinal obtido do aparelho receptor e i representa o último termo do vetor, portanto, os padrões podem ser obtidos com base no canal, como em

$$h(n) = s(n) - 0.5s(n-1) + s(n-2) + 0.5s(n-3) \quad (14)$$

Aqui, $h(n)$ é produzido um valor estimado de saída no sistema receptor, considerando a influência dos valores anteriormente transmitidos no canal, o que caracteriza a interferência inter-simbólica.

De forma a exemplificar o processo, considere a sequência de bits $X = \{1010\}$, que pode ser representado por $X = \{x_1, x_2, x_3, x_4\}$, e os coeficientes da função de transferência do canal, o produto dos bits com os coeficientes do canal 14, formam um processo de convolução, como em

$$\begin{aligned} h_1 &= x_1 s_1 \\ h_2 &= x_2 s_1 + x_1 s_2 \\ h_3 &= x_3 s_1 + x_2 s_2 + x_1 s_3 \\ h_4 &= x_4 s_1 + x_3 s_2 + x_2 s_3 + x_1 s_4 \end{aligned}$$

ou em termos práticos,

$$\begin{aligned} h_1 &= [1] \times 1 \\ h_2 &= [0] \times 1 - [1] \times 0.5 \\ h_3 &= [1] \times 1 - [0] \times 0.5 + [1] \times 1 \\ h_4 &= [0] \times 1 - [1] \times 0.5 + [0] \times 1 + [1] \times 0.5 \end{aligned}$$

onde cada produto interno representa um atributo do padrão que vai entrar na rede neural, e a variável alvo é o próprio bit que vai ser treinado. Essa padronização foi adotada para se adequar à entrada e saída do classificador.

Os valores binários são submetidos ao mesmo canal, mas com diferentes ambientes ruidosos a partir da relação sinal-ruído (SNR). SNR é uma medida usada nos sistemas de comunicação para comparar a potência do sinal e a potência do ruído como em (15), no entanto, os valores do sinal geralmente são expressos usando a escala de decibéis logaritmos como em (16)

$$SNR = \frac{P_{sinal}}{P_{ruído}}. \quad (15)$$

$$SNR_{dB} = 10 \log_{10} \left(\frac{P_{sinal}}{P_{ruído}} \right). \quad (16)$$

O conjunto de dados são submetidos a uma transmissão que passa por um dos oito ambientes ruidosos com base na relação sinal-ruído, sendo 2, 4, 6, 8, 10, 12, 14 e 16 em dB os SNR. Estes oito ambientes vão formar os conjuntos de dados que são testados em experimentos, essa etapa ocorre depois que o sinal passa pelo canal de comunicação.

Para detectar o erro, usamos como regra principal o conceito apresentado na definição da regra de decisão ótima, principalmente a Equação 13, onde um padrão (ou bit) deve ser rejeitado se o modelo compreender que existe uma dúvida em classificá-lo. Esse processo não necessita verificar bits redundantes, pois o simples processo de rejeição, já é considerado como um bit detectado que pode estar errado.

Os padrões são classificados usando o conceito de 80% dos padrões para treinamento e 20% dos padrões para teste. Na Tabela I é mostrado o balanceamento dos dados. Para transmitir a informação com integridade, os bits são codificados utilizando a implementação de 8 bits de dados por pacote e mais o bit de paridade, sendo assim, o total de 9 bits, padronização utilizada para se adequar ao detector de erro SPC. O que significa que com a implementação de bit paridade, os 2000 bits de dados e 250 bits de redundância, total de 2250 bits, se tornam 250 pacotes de dados.

Nas simulações computacionais é utilizada a validação cruzada de $k - Folds$ para encontrar os melhores parâmetros dos classificadores. O método separa o conjunto de treinamento em k partições e para cada configuração de parâmetro, um classificador é treinado com uma das combinações possíveis. O erro médio de todas as partes de k são calculadas para determinar a melhor configuração de parâmetros. Neste trabalho, para não aumentar o tempo de treinamento do classificador são utilizados 5 $Folds$ e 20 realizações para computar os resultados.

Tabela I
INFORMAÇÕES SOBRE O CONJUNTO DE DADOS

#	Treinamento		Teste		Total
Classe	Quantidade	(%)	Quantidade	(%)	(%)
0	4006	50.07	998	49.90	50.04
1	3994	49.93	1002	50.10	49.96
Total	8000	100.0	2000	100.0	100.0

Por último, o ambiente de execução de treinamento e testes foram realizados em um computador com um processador i5, uma frequência de 3,4 GHz e 16GB de memória. Ubuntu 20.04 LTS de 64 bits. Todos os algoritmos foram implementados na linguagem Matlab.

VII. RESULTADOS

Na Figura 3.(a) é apresentado o gráfico de dispersão dos padrões, variando o SNR de 2 a 16 dB para o conjunto de dados do Artificial. Os atributos dos padrões estão normalizados entre 0 a 1. Os padrões em cor vermelha (o) podem ser considerados como sendo padrões de classe (zero) e os padrões de cor azul (+) podem ser considerados como sendo padrões de classe (um). Os eixos X1 e X2 são atributos de padrões do Artificial. Na Figura 3.(b) é apresentada as curvas de Acurácia-Rejeição (A-R) para os algoritmos de classificação que são comparados. Observe em todos os métodos um comportamento que aumenta a taxa de acerto à medida que a taxa de rejeição também aumenta. As curvas (A-R) mos-

tram um comportamento semelhante entre os algoritmos de classificação de padrões.

Na Tabela II são mostrados resultados quantitativos para acurácia e desvio padrão das simulações da Figura 3. Os valores de acurácia e rejeição levam em consideração a soma de todos os custos de rejeição aplicados no modelo. Para SNR de 2 dB o MLP, RBF e ELM com opção de rejeição alcançaram respectivamente taxas de acerto de 84.49%, 84.49% e 86.93%. Considerando a classificação convencional (sem a opção de rejeição) a taxa de acerto é respectivamente a 72.45%, 72.56% e 73.35% para MLP, RBF e ELM. Para SNR de 16 dB os algoritmos testados alcançam 100% de acurácia.

Tabela II
ACURÁCIA E DESVIO PADRÃO DOS ALGORITMOS DE APRENDIZADO DE MÁQUINA PARA O ARTIFICIAL (%).

SNR (dB)	MLP	RBF	ELM
2	84.94±08.82	84.49±07.64	86.93±13.64
4	87.37±08.07	87.49±08.30	89.79±10.09
6	90.57±05.01	90.59±04.75	90.62±06.78
8	93.23±02.52	93.56±02.19	93.67±03.03
10	96.45±02.02	97.21±01.78	97.31±02.64
12	99.07±00.51	98.95±00.49	98.96±00.63
14	99.53±00.07	99.54±00.10	99.57±00.21
16	100.0±00.00	100.0±00.00	100.0±00.00

Na Tabela III são apresentados os resultados considerando a acurácia e rejeição. À medida que os valores de SNRs aumentam, as métricas de acerto também aumentam e a taxa de rejeição diminui. Para SNR de 2 dB o MLP, RBF e ELM possuem respectivamente taxa de rejeição média em 55.90%, 56.93% e 47.01%. Para SNR de 16 dB o MLP e ELM possuem taxa de rejeição média em 00.00%, significando que neste ambiente ruidoso não tem mais padrões rejeitados. O RBF ainda utiliza o método de opção de rejeição, pois a taxa de rejeição média para SNR de 16 dB está em 00.01%. Embora não seja o objetivo deste artigo, o ELM apresenta os melhores resultados com a opção de rejeição.

Tabela III
ACURÁCIA E MÉDIA DE REJEIÇÃO DOS ALGORITMOS DE APRENDIZADO DE MÁQUINA PARA O ARTIFICIAL (%).

SNR (dB)	MLP		RBF		ELM	
	A	R	A	R	A	R
2	84.36	55.90	84.49	56.93	86.93	47.01
4	87.37	44.77	87.49	46.85	89.79	39.42
6	90.57	19.63	90.59	22.30	90.62	30.80
8	93.23	20.83	93.56	29.87	93.67	26.71
10	96.45	18.81	97.21	25.42	97.31	24.48
12	99.07	06.74	98.95	10.10	98.96	12.71
14	99.53	01.41	99.54	01.51	99.57	03.06
16	100.0	00.00	100.0	00.01	100.0	00.00

A Tabela IV mostra os resultados para os detectores de erros SPC e com rejeição para uma realização. Para SNR em 2 dB (ambiente muito ruidoso), 237 pacotes chegaram errados na recepção (94,80%), o SPC detectou somente 134 pacotes como sendo errados (53,60%), enquanto que o método de rede neural

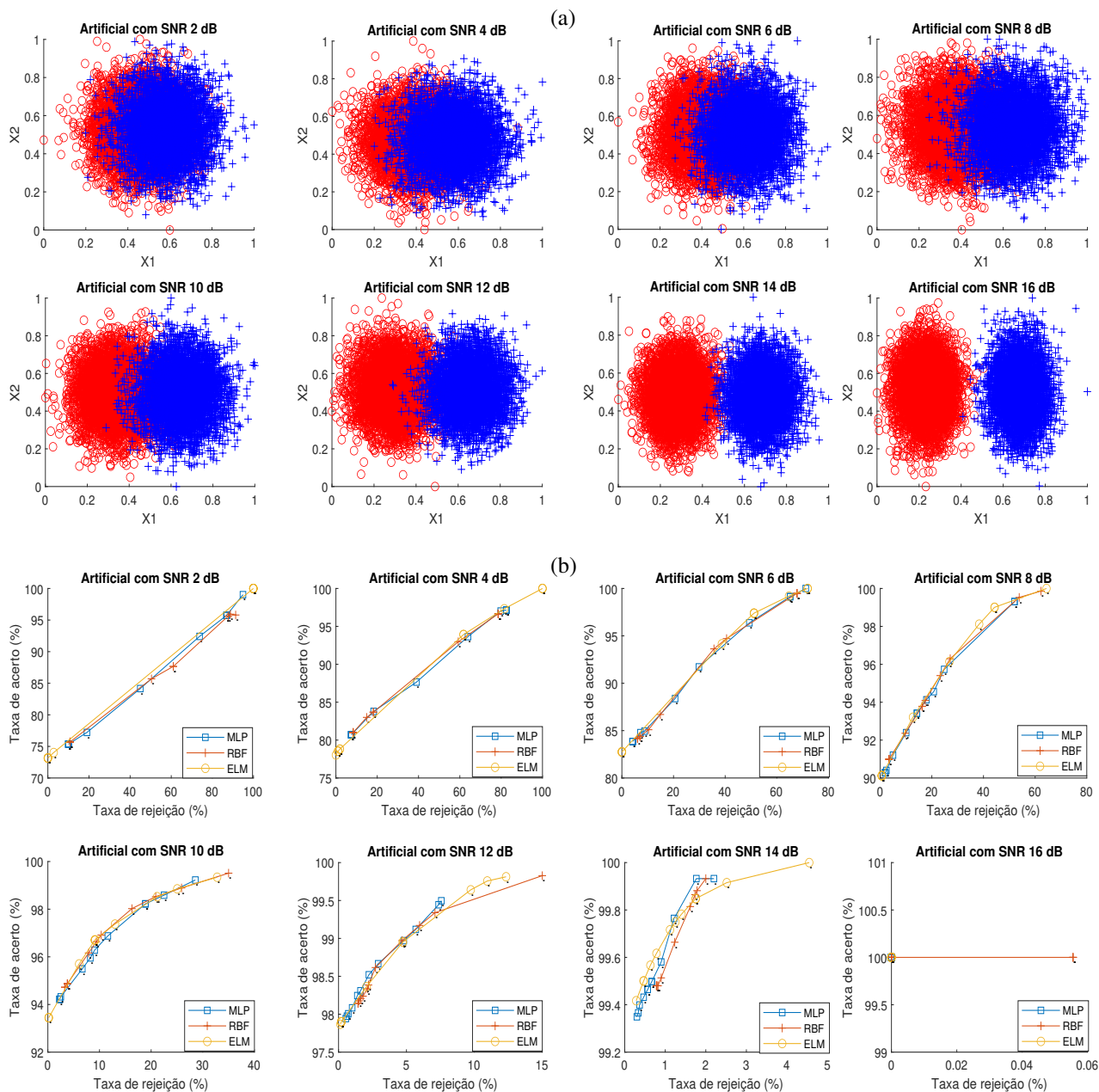


Figura 3. Algoritmos de aprendizado de máquina com opção de rejeição para o Artificial com SNR variando de 2 a 16 dB. (a) Superfície dos padrões do Artificial. (b) Curvas A-R do Artificial.

com opção de rejeição informou 242 (96,80%). Para SNR em 8 dB, 151 (60,40%) dos pacotes chegaram errados na recepção, o SPC detectou 101 (40,40%) pacotes como sendo errados, enquanto que o método com rejeição informou 158 (63,20%). O método com opção de rejeição é mais conservador, tem a preferência de rejeitar um pacote duvidoso que aceitar e errar, por isso que em alguns casos a detecção de pacotes errados é maior que a quantidade de realmente de pacotes errados. O método com rejeição é mais convergente com a situação de pacotes errados. Para SNR em 14 dB, 11 pacotes chegam errados e o com o método de rejeição detecta erros em 11

pacotes. O detector SPC não mostra eficiência para ambientes muito ruidosos.

VIII. CONCLUSÃO

Os classificadores com opção de rejeição propostos neste trabalho, realizam a tarefa de equalização de canais de comunicação e detectam erros em canais de comunicação. Como é de se esperar, o desempenho de classificação, em geral, se torna mais eficiente à medida que são rejeitados mais padrões, os quais também tem maior probabilidade de serem mal rotulados. A técnica proposta consegue detectar pacotes

Tabela IV
DESEMPENHO DE DETECÇÃO DE ERRO COM SPC E COM REJEIÇÃO

SNR (dB)	Pacotes recebidos (250) Pacotes errados	Detecção de erro	
		SPC	Rejeição
2	237	134	242
4	211	113	216
6	202	123	210
8	151	101	158
10	94	75	96
12	58	51	57
14	11	10	11
16	2	2	2

com erros através do processo de rejeição de padrões e, assim, a aplicabilidade da técnica de opção de rejeição na área de comunicação de dados é validada. Vale ressaltar ainda que o método proposto apresenta melhores resultados para detecção de erros quando comparado ao estabelecido SPC.

REFERÊNCIAS

- [1] Z. Z. Wengui. Error detection and control of IIoT network based on CRC algorithm. *Computer Communications*, v. 153, p. 390-396, 2020.
- [2] B. Kavita, Y. Ram, S. Shrivastava. Channel equalization using neural networks: A review. *IEEE transactions on systems, man, and cybernetics, Part C (Applications and Reviews)*, v. 40, n. 3, p. 352-357, 2010.
- [3] D. Carrera, C. Vargas-Rosales, M. Noe, L. Azpilicueta. Comparative study of artificial neural network based channel equalization methods for mmWave communications. *IEEE Access*, v. 9, p. 41678-41687, 2021.
- [4] Z. Qingyi, Z. Fan, Y. Chuanchuan. Adann: Adaptive neural network-based equalizer via online semi-supervised learning. *Journal of Lightwave Technology*, v. 38, n. 16, p. 4315-4324, 2020.
- [5] C. K. Chow. On optimum recognition error and reject tradeoff. *IEEE Transactions on information theory*, v. 16, n. 1, p. 41-46, 1970.
- [6] C. K. Chow. An optimum character recognition system using decision functions. *IRE Transactions on Electronic Computers*, n. 4, p. 247-254, 1957.
- [7] A. R. Rocha Neto, S. Ricardo, G. A. Barreto, J. S. Cardoso. Diagnostic of pathology on the vertebral column with embedded reject option. In: *Iberian Conference on Pattern Recognition and Image Analysis*. Springer, Berlin, Heidelberg, 2011. p. 588-595.
- [8] C. F. Bioucas-Dias, J. Castro, C. A. Ozolek, J. A. K. Jelena. Classification with reject option using contextual information. In: *2013 IEEE 10th International Symposium on Biomedical Imaging*. IEEE, 2013. p. 1340-1343.
- [9] R. Sousa, B. Mora, C. S. Jaime. An ordinal data method for the classification with reject option. In: *2009 International Conference on Machine Learning and Applications*. IEEE, 2009. p. 746-750.
- [10] M. D. Rocha, L. S. Gomes, J. P. P. Neto, A. R. Rocha Neto. Classification with reject option for software defect prediction. *Applied Soft Computing*, v. 49, p. 1085-1093, 2016.
- [11] I. Pillai, G. Fumera, F. Roli. Multi-label classification with a reject option. *Pattern Recognition*, v. 46, n. 8, p. 2256-2266, 2013.
- [12] A. Ahmadi, S. Omatu, T. Fujinaka, T. Kosaka. Improvement of reliability in banknote classification using reject option and local PCA. *Information Sciences*, v. 168, n. 1-4, p. 277-293, 2004.
- [13] M. Garima, S. Amandeep S. *Adaptive equalization algorithms: An overview*. *International Journal of Advanced Computer Science and Applications*, v. 2, n. 3, p. 62-67, 2011.
- [14] S. T. F. Barbosa. *Equalização neural aplicada a sistemas com modulação bidimensional em fibra óptica*. Dissertação de Mestrado. Universidade Federal do Rio Grande do Norte. n.º p. 81, 2014.
- [15] K. Sen-Maw e G. Woon-Seng. *Digital signal processors: architectures, implementations, and applications*. Prentice Hall, 2005.
- [16] T. Li; J. Jean. *Digital signal processing: fundamentals and applications*. Academic Press, 2013.
- [17] V. A. Pedroni. *Eletrônica digital moderna e VHDL*. Editora Campus, 2010.
- [18] A. Grami. Chapter 3-signals, systems, and spectral analysis. *Introduction to Digital Communications*, p. 41-150, 2016.
- [19] F. Aref, M. Farshid, A. Masood. Implementation of multilayer perceptron (MLP) and radial basis function (RBF) neural networks to predict solution gas-oil ratio of crude oil systems. *Petroleum*, v. 6, n. 1, p. 80-91, 2020.
- [20] D. Rumelhart, D. Hinton, W. Ronald. Learning representations by back-propagating errors. *nature*, v. 323, n. 6088, p. 533-536, 1986.
- [21] D. Broomhead, D. Lowe. Multivariable functional interpolation and adaptive networks, complex systems, vol. 2. 1988.
- [22] K. Bizona, G. Continillo, S. Lombardia, E. Mancaruso, B. Vaglieco. ANN-based Virtual Sensor for On-line Prediction of In-cylinder Pressure in a Diesel Engine. In: *Computer Aided Chemical Engineering*. Elsevier, 2014. p. 763-768.
- [23] H. Xiangdong, A. Lapedes. Nonlinear modeling and prediction by successive approximation using radial basis functions. *Physica D: Nonlinear Phenomena*, v. 70, n. 3, p. 289-301, 1994.
- [24] AP Braga. *Redes neurais artificiais: teoria e aplicações*. Livros Técnicos e Científicos, 2000.
- [25] H. Guang-Bin, Z. Qin-Yu, S. Chee-Kheong. Extreme learning machine: theory and applications. *Neurocomputing*, v. 70, n. 1-3, p. 489-501, 2006.
- [26] G. Fumera, F. Roli. Support vector machines with embedded reject option. In: *International Workshop on Support Vector Machines*. Springer, Berlin, Heidelberg, 2002. p. 68-82.
- [27] A. A. Bounsiar, E. Grall, P. Beausery. A kernel based rejection method for supervised classification. *International Journal of Computational Intelligence*, v. 3, n. 4, p. 312-321, 2006.