

As Redes Neurais de Hopfield e Multi-Layer Perceptrons Formando uma Arquitetura Neural Híbrida (MLP+H) com Características Próprias

Clayton Silva Oliveira, Emílio Del Moral Hernandez
Escola Politécnica da Universidade de São Paulo
Departamento de Engenharia de Sistemas Eletrônicos
Cidade Universitária – Av. Prof. Luciano Gualberto
São Paulo, SP, CEP 05508-900 – Brasil
E-mails: clayton@uol.com.br, edmoral@usp.br

Abstract

This article has the objective of presenting the first results of the project of one hybrid neural network composed by the junction of two different types of intelligent computational systems: the Hopfield neural network and the Multi-Layer Perceptron. In this sense, it was possible to develop one hybrid neural system that presents better performance, in some cases, than the two cited types of neural networks. For example, the immunity to external noises added to the input patterns was improved. Additionally, this hybrid system has the ability of dealing with different types of data when compared with classical architectures. For example, while the Hopfield network deals with binary patterns, and the MLP produces one function that maps inputs to outputs by its training, dealing with information that have always analog characteristics, the hybrid network MLP+H works with analog inputs and digital outputs. As it will be discussed, the hybrid neural architecture developed in the project here presented has a good flexibility in various applications, as anti-noise filters, what is actually being studied. Also important in this paper is the discussion of concepts and methodologies of evaluation and characterization of performance of the considered architectures based on the confusion matrix concept, which allows a deeper analysis of experimental results.

1 Introdução

As redes neurais do tipo Multi-Layer Perceptron (MLP) e Hopfield possuem características únicas que permitem a resolução de problemas específicos dentro do universo de Inteligência Artificial [1]. Porém, problemas que requerem necessidades específicas e não encontradas em somente um tipo de arquitetura justificam o desenvolvimento de arquiteturas híbridas.

Dessa forma, como será visto, a arquitetura híbrida aqui analisada reúne características fundamentais

daqueles dois tipos de redes neurais, para formar outras distintas e únicas, que resolvam problemas específicos com melhores desempenhos para um mesmo tipo solucionável por alguma das duas arquiteturas isoladamente, e eventualmente com menores tempos computacionais para o treinamento da rede.

2 Discussão de Características e de algumas limitações das redes de Hopfield e MLP

A rede de Hopfield é definida como sendo uma memória associativa, pois armazenando padrões (atratores), ela pode recuperá-los posteriormente ao seu treinamento a partir de uma entrada, que, geralmente, é um dos padrões previamente armazenados, porém disponível apenas em forma parcial ou distorcida por ruído [2]. É também conhecida como uma memória auto-associativa, já que, diferentemente dos perceptrons, por exemplo, que podem armazenar pares de padrões (entrada-saída) de naturezas diferentes (como por exemplo as dimensões da entrada e da saída distintas, ou entradas digitais e saídas analógicas), esse tipo de arquitetura só relaciona dados de mesma natureza. A figura 1 exemplifica uma rede de Hopfield.

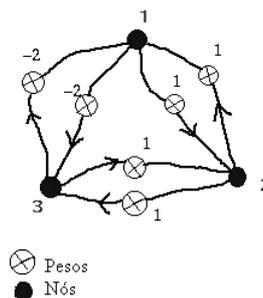


Figura 1: Um exemplo da arquitetura de Hopfield, com 3 nós [3]. Percebe-se a ligação entre todos os nós da rede, de forma que qualquer nó se comunique com os restantes. Adicionalmente, verifica-se que o peso da ligação entre um nó e outro é o mesmo desse outro para o um. A forma de ligação entre os neurônios de uma rede neural é um dos fatores que definem suas características.

¹ Acknowledgments to the University of São Paulo (USP) and Foundation for the Support of Research in the State of São Paulo (FAPESP) for the financial support to this work.

Após seu treinamento, uma vez fornecido à rede um vetor binário de dimensão N (sendo N o número de nós da rede), esta deve ser capaz de recuperar um dos vetores (padrões) armazenados.

De forma geral, pode-se definir algumas características e limitações da rede de Hopfield particularmente relevantes para as discussões que se seguem:

- Ocorrência de respostas erradamente armazenadas, ou que geralmente não convêm (padrões espúrios e padrões espelhos, como exemplos);
- Relação apenas entre entradas e saídas de mesma natureza (entradas digitais e saídas digitais, e com mesmo número de bits em ambas);
- Após o treinamento, desconsideradas as respostas que não convêm, a rede de Hopfield fornece apenas respostas pré-armazenadas para todas as entradas, ou seja, dada uma entrada qualquer, a rede indicará como vetor de saída um dos padrões de treinamento armazenados que mais se aproxime daquela entrada. Dessa forma, essa arquitetura neural tem a característica de consertar *erros* das entradas, como bits trocados (por exemplo, 1 por 0).

Já a arquitetura Multi-Layer Perceptron (MLP) se caracteriza, principalmente, pelo tipo de resposta que cada nó pode fornecer: diferentemente da rede de Hopfield, que pode unicamente trabalhar com entradas e respostas *binárias*, o MLP pode possuir entradas e saídas *analógicas* [4]. Um exemplo de sua arquitetura é dada na figura 2.

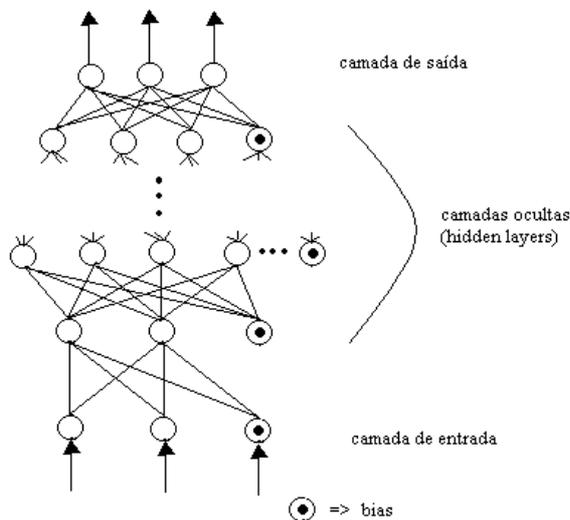


Figura 2: A estrutura do Multi-Layer Perceptron. Como se pode notar, existe uma entrada e uma saída da rede bem definidas, diferentemente da rede neural de Hopfield (figura 1). A ligação de seus nós em camadas garante características distintas daquelas encontradas em outras redes neurais, como a de Hopfield, por exemplo.

Para a aprendizagem da rede MLP, é fornecido um conjunto de treinamento formado por pares entrada-saídas, dos quais queremos que a rede aprenda a relação existente entre estes pares.

Após a fase de aprendizagem, o MLP será capaz de, dado qualquer vetor como entrada, fornecer como resposta o vetor *mais adequado*, de acordo com a relação entre os padrões de entrada e os padrões de saída, aprendida previamente através do conjunto de treinamento.

Pode-se listar algumas características da rede do tipo MLP também relevantes aos temas abordados neste artigo:

- Pode trabalhar com conjuntos de treinamento de naturezas diferentes: por exemplo, entradas digitais com dimensão N_E , saídas analógicas com dimensão N_S ;
- Depois do treinamento, dada uma entrada, o MLP fornecerá como saída um vetor qualquer que mais se adeque àquela entrada de acordo com a função aprendida (que relaciona entradas e saídas fornecidas) durante a fase de aprendizagem. Dessa forma, o MLP se caracteriza por ser um *interpolador e extrapolador*;
- Por ser uma arquitetura neural flexível, capaz de resolver muitos problemas, como divisões dos dados de treinamento em classes sem separabilidade linear (o problema XOR é o exemplo clássico [4]), seu treinamento é bem mais demorado quando comparado com o treinamento da rede de Hopfield, o que representa grandes tempos computacionais envolvidos.

Com a apresentação das principais características dos dois tipos de arquiteturas neurais que serão componentes da rede neural híbrida aqui apresentada, ilustra-se também alguns conceitos chaves da estrutura e funcionamento da MLP+H.

3 A rede híbrida proposta: Multi-Layer Perceptron + Hopfield (MLP+H) – características funcionais e estruturais da rede MLP+H

Como veremos a seguir, a arquitetura híbrida desenvolvida e aqui apresentada pode relacionar pares entrada-saída nos quais a entrada pode tanto ser analógica quanto discreta, e a saída é sempre de natureza discreta, não importando a natureza da entrada; outra característica é a possibilidade de diferentes dimensões para as entradas e saídas. Dessa forma, ela possui uma característica que nenhuma das duas redes neurais clássicas aqui discutidas possui individualmente: relacionar *entradas analógicas* com *saídas puramente discretas* com mecanismo de correção de erros, já que o MLP, por exemplo, possui características que permitem entradas e saídas analógicas sem nenhum mecanismo de correção de erros [2, 4, 5], como os atratores na rede de Hopfield, por exemplo (Seção 2).

Estruturalmente, a arquitetura híbrida MLP+H é definida como representado na figura 3.

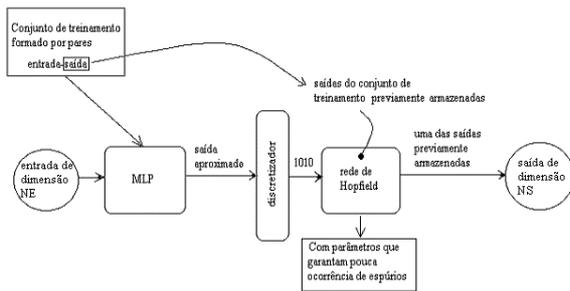


Figura 3: A estrutura da arquitetura híbrida MLP+H. O encadeamento das redes neurais MLP e Hopfield garante a essa arquitetura híbrida características distintas quando comparada com suas arquiteturas componentes consideradas isoladamente. Por exemplo, pode-se ter entradas exclusivamente analógicas de dimensão N_E com saídas exclusivamente digitais de dimensão N_S .

Como podemos observar, a dimensão da entrada (N_E) pode ser diferente da dimensão da saída (N_S), o que não é possível para a rede de Hopfield em separado. Dessa forma, a rede híbrida MLP+H estruturalmente é fundamentada como o encadeamento de uma MLP seguida de uma rede de Hopfield.

Durante a concepção inicial da arquitetura híbrida MLP+H, uma das opções consideradas inicialmente foi o acoplamento direto entre a saída do MLP e da rede de Hopfield. Outra possibilidade considerada posteriormente foi a discretização (binarização, para ser mais exato) da saída do MLP, antes do fornecimento da entrada à rede de Hopfield. Como observaremos na parte experimental, a performance dessa rede sem o discretizador era pior do que a performance de um MLP com sua saída discretizada. A explicação do porquê a inserção de um discretizador entre o MLP e a rede de Hopfield melhorar a performance da arquitetura híbrida, como será mostrado na Seção 5, é a seguinte: a rede de Hopfield recebe entradas *binárias*, e não *analógicas* como na idéia original, o que faz com que a rede de Hopfield trabalhe com entradas mais próximas de sua natureza discreta.

A funcionalidade da recuperação dos padrões armazenados é caracterizada dessa forma:

- dada uma entrada ao MLP (ruidosa ou não), com N_E dimensões, esta produzirá uma resposta de acordo com as interconexões entre os neurônios e as camadas do Multi-Layer Perceptron definidas no treinamento;
- essa saída, depois de *binarizada*, será dada como entrada à rede de Hopfield, que fará iterações até que uma saída previamente armazenada seja dada como sendo um *estado estável*; esse estado será a saída da rede híbrida MLP+H, com N_S bits, que

pode ser um dos padrões armazenados ou não (pode ser um padrão espúrio ou espelho também).

Funcionalmente, essa arquitetura permite, por exemplo, que as entradas e saídas sejam diferentes quanto ao domínio, ou seja, as entradas podem ser discretas ou analógicas, enquanto que as saídas sempre serão discretas (pois a rede de Hopfield está na saída da MLP+H), o que não é possível para o MLP em separado, já que esse tipo de arquitetura neural, estritamente falando, lida com entradas e saídas que sempre tem algum caráter analógico.

Sobre o algoritmo de treinamento, ele é definido da seguinte forma:

- previamente, as saídas (sempre binárias) do conjunto de treinamento são armazenadas na rede de Hopfield da MLP+H;
- treina-se o MLP com os pares do conjunto de treinamento até que a rede de Hopfield seja capaz de reconhecer as saídas fornecidas pelo MLP como sendo as previamente armazenadas.

Do exposto acima, pode-se esperar que a arquitetura MLP+H possua um treinamento mais rápido do que o MLP sendo treinado solitariamente: o MLP não precisa ser treinado até que seu erro RMS seja menor do que um valor requerido; basta ser treinado até o ponto em que suas saídas possam ser reconhecidas pela rede de Hopfield. Como veremos na Seção 5, seção em que resultados experimentais concretos são discutidos, de fato a estratégia proposta permite uma boa redução no tempo de aprendizagem da rede neural. Um outro fato relevante é que mesmo que o MLP não consiga ser treinado para aquele conjunto fornecido para a aprendizagem, é possível que a rede híbrida, em poucas iterações, tenha seu treinamento concretizado².

Essa rede híbrida pode conseguir melhores performances em relação a estados espúrios quando comparada com a arquitetura de Hopfield em separado, pois o MLP (com um discretizador na saída) fornece como entrada à rede de Hopfield vetores binários suficientemente próximos aos previamente armazenados, o que acaba aumentando a capacidade dessa rede de reconhecê-los como um dos previamente armazenados, já que uma das condições do fim de treinamento é que a rede de Hopfield esteja reconhecendo todos os padrões que o MLP esteja fornecendo a ela.

Para melhor ilustração e consolidação dessas idéias, em seção à frente vamos mostrar os resultados de diversos experimentos feitos com a rede híbrida MLP+H. Antes porém, na Seção 4 que segue vamos falar um pouco de uma ferramenta de caracterização importante para a análise de tais experimentos.

² Em teste preliminar realizado e não apresentado neste artigo.

4 O conceito de matriz de confusão (confusion matrix) para a medição da performance de arquiteturas neurais de natureza discreta

Com o decorrer do desenvolvimento de nossos estudos sobre a arquitetura híbrida MLP+H, os testes exigidos para a caracterização dessa ficavam cada vez mais complexos e em maiores quantidades, pois como as características da arquitetura híbrida eram desconhecidas, os testes feitos com essa rede neural eram de vital importância para a sua caracterização e levantamento de sua performance frente à entrada de versões ruidosas de padrões armazenados.

Do exposto acima, pode-se deduzir que a verificação dos acertos da rede neural por meio da análise de tabelas que relatam para cada entrada a saída fornecida pela rede neural, seria extremamente trabalhosa e não permitiria testes mais exaustivos; dessa forma, o conceito de *matriz de confusão* (*confusion matrix*) foi de extrema importância para os testes exigidos por nossos estudos.

O conceito de *matriz de confusão* é muito utilizado em Telecomunicações, pois tem a capacidade de reunir, em uma só tabela (matriz), como um certo canal de comunicação se comporta frente a diversos padrões que por ele passam no que tange à conservação das características próprias desses padrões ao longo do canal de comunicação, ou seja, dado um padrão P como entrada ao canal, se este chegará na saída com as mesmas características, ou se será “confundido” pelo receptor de mensagens como sendo um dos demais padrões. Dessa forma, um bom canal de comunicação consegue conservar o padrão desde a entrada até a saída, ou melhor, não introduz componentes de ruído ao sinal (padrão) em questão. No nosso contexto, a matriz de confusão fornece a seguinte informação: dado o vetor P de entrada, quais são as probabilidades de que ele seja identificado na saída como o próprio vetor P , ou como um dos outros vetores armazenados.

Nesse caso, precisa-se saber como uma rede neural se comporta para cada versão (ruidosa ou não) dos padrões previamente armazenados, ou seja, dado uma versão ruidosa de um certo padrão P , queremos saber quais são as probabilidades de que este seja reconhecido como o padrão P , como um outro padrão armazenado, ou ainda como um vetor espúrio. Assim, para esse caso, a matriz de confusão é dada pela seguinte estrutura da Tabela 1:

Tabela 1: A estrutura de uma matriz de confusão

		Padrões dados como <i>saída</i> da arquitetura neural		
		Padrão 1	Padrão 2	... Padrão M
n versões ruidosas do padrão i dadas como <i>entrada</i>	Padrão 1	X_1 reconhecidos como padrão 1 (acertos)	Y_1 reconhecidos como padrão 2 (erros)	Z_1 reconhecidos como padrão M (erros)
	Padrão 2	X_2 reconhecidos como padrão 1 (erros)	Y_2 reconhecidos como padrão 2 (acertos)	Z_2 reconhecidos como padrão M (erros)
	... Padrão M	X_3 reconhecidos como padrão 1 (erros)	Y_3 reconhecidos como padrão 2 (erros)	Z_3 reconhecidos como padrão M (acertos)

Explicando melhor a matriz, tem-se:

- tendo-se a primeira linha referente às n entradas ruidosas do padrão 1, cada coluna se refere ao padrão de saída dado pela rede, ou seja, a primeira coluna (elemento₁₁ da matriz) fornece o número X_1 de *acertos* da rede para esses n testes, enquanto que as demais colunas fornecem $Y_1 + Z_1$ *erros*, ou melhor, $Y_1 + Z_1$ versões ruidosas do padrão 1 foram reconhecidas de forma errada como padrão 2 ou padrão k ;
- $X_i + Y_i + \dots + Z_i = n$, sendo n o número de versões ruidosas do padrão i ($i = 1, 2, \dots, k$) dadas como entrada à rede;
- caso $X_i + Y_i + \dots + Z_i \neq n$, então $[n - (X_i + Y_i + \dots + Z_i)] =$ número de vetores espúrios fornecidos como resposta pela rede.

Dessa forma, pode-se perceber que a diagonal principal da matriz revela o quanto a rede é robusta a ruídos, já que a *soma* dos elementos da diagonal principal fornece o número total de acertos da rede neural testada.

Uma forma interessante de se apresentar a matriz de confusão é em termos de porcentagem de acertos/erros, ao invés dos números absolutos, já que por ser uma medida relativa, é de mais fácil comparação entre matrizes relativas a diferentes redes neurais.

Tendo-se em mente todo o conceito de matriz de confusão apresentado, pode-se inferir que uma forma concisa de se visualizar a performance da arquitetura neural testada é a seguinte: basta calcular a *média* da diagonal principal dessa matriz, e portanto tem-se uma *média* da performance, ou número de acertos da rede; dessa forma, quanto maior esse número, maior será a imunidade da rede diante de ruídos.

Para melhor ilustração e exemplificação desse conceito, apresentamos a seguinte matriz de confusão na Tabela 2, resultado de um teste com uma rede neural:

Tabela 2: Um exemplo de aplicação do conceito de matriz de confusão em um teste real

		Padrões dados como <i>saída</i> da arquitetura neural		
		Padrão 1	Padrão 2	Padrão 3
n versões ruidosas do padrão i dadas como <i>entrada</i>	Padrão 1	<u>60%</u>	20%	10%
	Padrão 2	10%	<u>80%</u>	10%
	Padrão 3	25%	30%	<u>45%</u>

Da 1ª. linha da matriz acima, pode-se concluir que 60% das n versões ruidosas do padrão 1 foram reconhecidas como padrão 1, 20% como padrão 2 e 10% como padrão k ; como $60\% + 20\% + 10\% = 90\%$, tem-se que houve a ocorrência de *10% de vetores espúrios*. Dessa forma, a média de acertos foi de $(60\% + 80\% + 45\%)/3 = 61,67\%$.

Assim, conseguimos adotar uma forma simples de se obter a média de acertos de uma rede neural para um certo conjunto de testes, e como consequência uma forma de se medir a performance dessa rede.

Como veremos, esse conceito de matriz de confusão foi muito importante para a caracterização da arquitetura MLP+H, por representar uma forma concisa e fácil de se obter um número que represente a performance (média de acertos) da rede; além disso, se desejássemos detalhes do comportamento da arquitetura, bastava analisar a matriz de confusão, o que é uma tarefa bem mais fácil do que se analisar cada resposta da rede para cada entrada.

5 Experimentos feitos com a arquitetura neural híbrida MLP+H

Essa seção tem o objetivo de apresentar os testes executados para o levantamento das características e performance da rede híbrida MLP+H frente à entrada de padrões ruidosos posteriormente ao seu treinamento, bem como para comparação de performances entre outras duas alternativas que resolvam o mesmo problema.

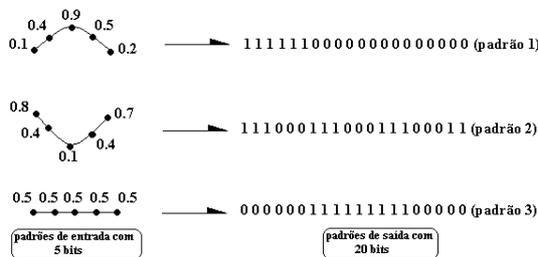


Figura 4: Conjunto de treinamento para o teste da arquitetura híbrida. Note a natureza analógica da entradas versus a natureza binária das saídas.

Primeiramente, fez-se um teste de caracterização, fornecendo-se à MLP+H o conjunto de treinamento representado na figura 4.

Os parâmetros da rede eram os seguintes:

- MLP de 3 camadas, com **6** ($5 + 1$ bias)/ **25** / **40** ($20 + 20$) nós;
- Hopfield de 40 nós, já que uma rede de Hopfield de 20 nós não possuiria *capacidade* aceitável para o armazenamento dos padrões de saída do conjunto de treinamento da figura 4 [6, 7]; dessa forma, os padrões de saída foram duplicados para se adaptarem à rede, já que possuíam 20 bits.

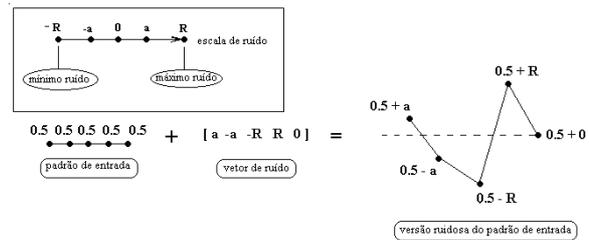


Figura 5: Ilustração da definição de ruído utilizada para medição da performance da rede MLP+H. Ao padrão de entrada, é somado um vetor de ruído formado por componentes que variam entre $-R$ e $+R$, distribuídos de forma uniformemente randômica.

É importante notar que obrigatoriamente a última camada do MLP deve possuir o mesmo número de nós da rede de Hopfield selecionada, já que essas estarão conectadas entre si.

Para a caracterização da arquitetura no que se refere à robustez ao ruído, definimos uma forma de *medir/insere* ruído nos padrões, criando-se assim versões ruidosas das entradas do conjunto de treinamento, no intuito de se medir a performance da rede híbrida em diferentes condições de entrada. Esse ruído foi definido pelo seu máximo/mínimo valor, ou seja, para um ruído máximo R e mínimo $-R$, sabe-se que aqueles padrões de entrada possuem em cada dimensão no máximo uma componente de ruído R ou no mínimo de $-R$. Essa componente de ruído é inserida nos padrões para a criação de testes apenas somando-se um vetor de ruído ao padrão do qual queremos criar uma versão ruidosa. A figura 5 ilustra o procedimento de geração de padrões de entrada com presença de ruído.

Com o conjunto de treinamento e os parâmetros da arquitetura definidos, foram dados às redes (MLP+H mais outras duas alternativas) 900 testes (versões ruidosas dos padrões), sendo que cada 300 testes se referiam a um só dos três padrões, gerados como ilustrado na figura 5.

Na figura 6, temos o gráfico que representa a performance (média da diagonal principal da matriz de confusão) das redes (MLP+H e MLP com a saída discretizada) *versus* máximo módulo do ruído $|R|$.

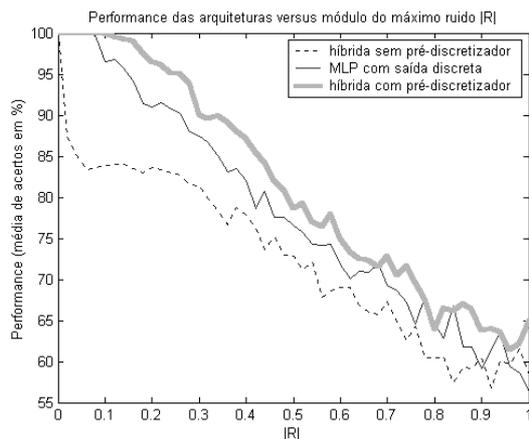


Figura 6: Performance das arquiteturas estudadas. O eixo das ordenadas se refere à performance obtida (média da diagonal principal da matriz de confusão); o eixo das abscissas se refere ao módulo da máxima componente de ruído. Como se pode observar, a performance da rede neural híbrida MLP+H (linha espessa) sempre é maior ou igual à performance de uma MLP com sua saída binarizada (linha fina), bem como de uma MLP+H sem a presença do discretizador entre suas componentes MLP e Hopfield (linha pontilhada). Note-se que quando o módulo do máximo ruído vale 1 (mínimo de 100% de ruído), a performance média da MLP+H fica em torno de 65% de acertos, enquanto que o MLP com a saída binarizada apresentou 55%.

Desse gráfico, pode-se verificar a validade dessa arquitetura híbrida para a implementação do conjunto de treinamento anteriormente apresentado e sua boa performance frente às entradas ruidosas, já que somente para $|R| > \sim 0,5$, a média de acertos fica abaixo de 80% dos testes apresentados; para reforçar ainda mais essa afirmação, é relevante dizer que os padrões de entrada possuem valores analógicos entre 0 e 1, e que um $|R| = 0,5$, significa 50% de ruído *no mínimo*, o que pode ser considerada uma alta componente de ruído. Além disso, como já mencionado, a performance da MLP+H foi melhor do que a do MLP com a saída discretizada.

Outro teste realizado visou contrastar a performance do MLP+H frente ao MLP, quanto ao número de iterações necessárias na fase de aprendizado. Para esse teste, o MLP possuía 3 camadas, cada qual com 4/15/20 nós, enquanto a híbrida MLP+H possuía um MLP semelhante a este, e uma rede de Hopfield de 20 nós. Assim, os números de iterações necessários para o treinamento de cada uma das redes (para um mesmo conjunto de treinamento) foram os seguintes: a rede híbrida MLP+H necessitou de 50 iterações; já o MLP, para um erro RMS de 0,07, necessitou de 163 iterações; para um erro de 0,05, 230; para um erro de 0,03, 434; e por fim, para um erro de 0,01, 1805 iterações.

Assim, como anteriormente comentado, a rede híbrida consegue “aprender” o mesmo conjunto de

treinamento apresentado ao MLP com *menos iterações*, pelos motivos já apresentados anteriormente.

Dessa forma, pode-se dizer que a rede MLP+H tem características únicas que podem resultar em melhores performances na resolução de problemas como o aqui apresentado, o que justifica seu uso no lugar de outras arquiteturas em problemas específicos, como é o caso relatado nesta seção.

6 Conclusões e perspectivas

Como pôde-se verificar, a rede neural híbrida MLP+H teve sua validade comprovada, sendo que muitas das vantagens inicialmente antecipadas quando da sua concepção efetivamente foram comprovadas experimentalmente: temos uma melhor performance quando da ocorrência de entradas ruidosas, quando a comparamos às outras arquiteturas estudadas; observa-se a necessidade de menores números de iterações de treinamento em comparação com o MLP em separado; por fim, a MLP+H possui características próprias e diferentes das redes MLP e Hopfield analisadas isoladamente, que permitem a solução de problemas que não seriam resolvidos nem por uma nem por outra, ou que em caso da solução ser possível, não apresentariam a mesma performance que a obtida com a MLP+H.

Sobre futuras perspectivas em aplicações práticas da arquitetura MLP+H, pode-se citar que a utilização dos conceitos e métodos desenvolvidos e aqui apresentados está atualmente sendo estudada no contexto de filtros de frequências [8] com características diferentes daquelas dos filtros analógicos e digitais convencionais, como a passagem com alta seletividade de frequências requeridas. Dessa forma, justifica-se teórica e experimentalmente a validade da rede híbrida MLP+H.

Referências

- [1] Haykin, Simon. Neural Networks: a comprehensive foundation. MacMillan Publishing Company, 1994.
- [2] Hopfield, J. J. “Neural Networks and Physical Systems with Emergent Collective Computational Abilities”, in Proc. National Academy of Sciences, USA, April 1982, Vol 79, pp.2554-58.
- [3] <http://www.kcl.ac.uk/neuronet/intro/index.html>
- [4] Rumelhart, D. E. and J. L. McClelland, eds, Parallel Distributed Processing: Explorations in the Microstructure of Cognition, V1, Cambridge, MA: MIT Press, 1986.
- [5] Kosko, B., “Bidirectional Associative Memories”, IEEE Trans. on Systems, Man, and Cybernetics, Vol. 18, pp.49-60. Jan-Feb, 1988.
- [6] Amit, D. J. Modeling Brain Function: the World of Attractor Neural Networks, New York: Cambridge University Press, 1989.
- [7] McEliece, R., E. Posner, E., and S. Venkatesh, “The Capacity of the Hopfield Associative Memory” Information Theory, IEEE Transactions on , Volume: 33 Issue: 4 , Jul 1987 pp. 461 –482
- [8] Oppenheim, Alan V., and R. W. Schaffer, Discrete-Time Signal Proc., Englewood Cliffs, NJ, Prentice Hall, 2000.