Anais do VII Simpósio de Pesquisa Operacional e Logística da Marinha (SPOLM 2004), Rio de Janeiro: Centro de Análises de Sistemas Navais, v.1, p.653-667, 2004

Redes Neurais no Auxilio Pedagógico aos Cursos de Formação Militar

Henrique Guilherme do Amaral Santos
Escola de Administração do Exército - ESAEX
Rua Território do Amapá, 455 Salvador/BA
guilhermewar@hotmail.com

Melisa Fernandes Maia
Escola de Administração do Exército - ESAEX
Rua Território do Amapá, 455 Salvador/BA
tenalmelisa@esaex.mil.br

Éldman de Oliveira Nunes
Escola de Administração do Exército - ESAEX
Rua Território do Amapá, 455 Salvador/BA
eldman@bol.com.br

Resumo

Este artigo tem por objetivo mostrar como o uso de técnicas computacionais de redes neurais e inteligência artificial pode auxiliar as atividades pedagógicas em escolas militares (e outros centros de ensino). Com o uso de teoria de redes neurais e de um programa desenvolvido no programa *Borland Delphi 7.0*, foi feita uma previsão da classificação de final de curso dos alunos da Escola de Administração do Exército – EsAEX utilizando dados dos mesmos, anteriores ao seu ingresso no Curso de Formação. Como resultado, observou-se um bom desempenho da rede utilizada que apresentou erro de teste de apenas 21,08% e que com um teste de hipótese mostrou-se estatisticamente significante, apresentando uma taxa real de acerto de 47,05%. A identificação da classe de desempenho de alunos em estabelecimentos militares pode ser extremamente interessante para os docentes e pedagogos já que permite que se preparem instruções mais adequadas e customizadas, de acordo com as necessidades do aluno.

Palavras-Chave: Redes Neurais; Formação Militar; Previsão de Classificação.

Abstract

This article's objective is to show how the use of neural network and artifitial computational intelligence techniques can help in pedagogic activities in military schools (and others educational centers). Adopting neural network theory and a *Borland Delphi 7.0* developed program, a prediction on the final classification for the Escola de Administração do Exército – EsAEx's students was made, based on some of their pre-course data. Therefore, a good performance for the neural network was observed with a test error rate of only 21,08%, which proved to be stathistically significant, and a real correct prevision rate of 47,05%. The military educational institutes' students' proper performance class identification can be extremally important for teachers and pedagogy professionals since it can provide an useful tool in the preparation of better customized and adequated classes, respecting each student's needs.

Keywords: Neural Network; Military Education; Performance Prevision.

1. INTRODUÇÃO

O presente trabalho tem por objetivo desenvolver um método alternativo para auxiliar, de forma pedagógica, nas diversas Escolas Militares do país a formação básica de militares. Haja vista a importância que esta tem para o futuro das Forças Armadas Brasileira, procura-se acrescentar mais um meio de orientação aos instrutores destas escolas, de forma que seja possível prever com certo grau de confiança o desempenho do indivíduo no curso, para 'que os instrutores possam ter melhores informações sobre os pontos fortes e os pontos a melhorar de cada aluno e, assim, desenvolver aulas e instruções mais apropriadas e eficientes.

Pretende-se utilizar o modelo de Redes Neurais Artificiais, descrito na seção 2, para classificar os alunos, que ingressam em cursos militares, dentro de uma classe de performance (classificação ao final do curso) de acordo com suas características pessoais e desempenho no concurso de admissão. E posteriormente, utilizando o mesmo modelo, fazer previsões das classificações finais dos alunos no momento que estes ingressam nos respectivos cursos.

Este artigo pretende desenvolver um novo método de analisar o perfil de alunos em um curso de formação, partindo de suas características de origem, anteriores ao ingresso e que permitam aos organizadores, instrutores e pedagogos desses cursos desenvolver métodos de ensino e avaliação mais personalizadas e individualizadas ao perfil de seus clientes-alunos. Dessa forma o aproveitamento e rendimento daqueles poderá apresentar um melhor conceito elevando assim o padrão dos recursos humanos desenvolvidos nas Forças Armadas Brasileira e que são de essencial importância para esta.

Na seção 2 apresenta-se o referencial teórico, discute-se as características básicas dos cursos de formação militar e explana-se sobre a teoria de redes neurais. Na 3 sobre os Materiais e métodos, ou seja, o aplicativo que foi utilizado no artigo. Finalizando as ultimas seções discutem os resultados encontrados na pesquisa. A seção 4 trata dos testes e resultados, a 5 das discussões e a 6 apresenta a conclusão.

2. REFERENCIAL TEÓRICO

2.1. A FORMAÇÃO MILITAR

A manutenção do Estado depende inexoravelmente dos recursos de ataque e defesa de suas Forças Armadas e do capital humano capaz de sustentar as decisões estratégicas que fortalecem sua integridade física, política e econômica. Dessa forma, toda Força Armada deseja possuir e construir recurso humano bem qualificado, treinado e bem equipado para cumprir com sua missão.

Hoje, a carreira militar tem se tornado uma opção concorrida para os jovens. Por outro lado, a vida militar apresenta características particulares e peculiares como dedicação intensiva, sujeição a preceitos rígidos de disciplina e hierarquia, disponibilidade intensa, preparo físico, restrições na participação de atividades políticas e sindicais, risco de vida, entre outros.

Além desses aspectos, é importante citar o componente social que envolve o militar, de forma que sua educação está fortemente relacionada com a criação de profissionais bem capazes de liderar e trabalhar em grupo, além da criação de um espírito de corpo e do respeito pelos superiores. A necessidade de atingir estes objetivos é de extrema preocupação e responsabilidade e são eles que irão determinar o nível de profissionalismo e eficiência das Forças Armadas Brasileira.

Portanto, é de extrema importância e muito bem vindos os métodos que buscam aprimorar as técnicas utilizadas nas escolas e cursos militares para a formação e instrução de seus alunos. Especialmente, aquelas associadas a técnicas computacionais.

Hoje, a tecnologia já é utilizada em atividades acessórias do ensino, como na administração mais eficiente de escolas e centros de preparação (com previsões mais adequadas de provisões, materiais e ração, com conseqüente economia) e na segurança desses locais com sistemas de identificação mais confiáveis e controlados.

Porém, pode-se ampliar este recurso para desde a construção de programas que auxiliam na classificação de alunos em níveis especiais e na adaptação do ensino a grupos particulares de indivíduos até a construção de programas de simulação de combate individual e de treinamento de condução de veículos e aeronaves de guerra.

2.2. REDES NEURAIS

Pode-se definir os modelos conexionistas de computação, também chamados de redes neurais artificiais (RNAs), redes neurais ou sistemas de processamento paralelo distribuído (PDP), como uma estrutura de processamento de informação distribuída e paralela capaz de mapear sistemas complexos, sem ter a necessidade de conhecer eventuais modelos matemáticos que descrevem seu comportamento.

Uma RNA é formada por unidades de processamento, comumente chamadas de nós, neurônios ou células, interconectadas por arcos unidirecionais, também chamados de ligações, conexões ou sinapses.

A maneira como os neurônios são arranjados na rede neural define a arquitetura da rede, a qual é estreitamente relacionada ao tipo de conhecimento por ela armazenado.

A arquitetura mais comum dispõe os neurônios em camadas, de modo que os neurônios de uma dada camada possam conectar-se unicamente aos da camada imediatamente anterior e emitem seu sinal exclusivamente aos neurônios da camada posterior, como mostrado na Figura 1.

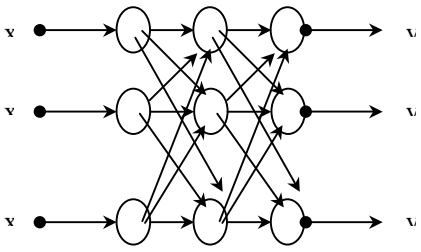

Figura 1 - Arquitetura Genérica de uma Rede Neural de Múltiplas Camadas

Os neurônios da camada entrada (*xi*) recebem os sinais provenientes do meio externo repassando-os à camada intermediária imediatamente posterior. Os neurônios das camadas intermediárias, ou escondidas, processam os sinais recebidos da camada anterior ponderados pelos pesos sinápticos, e, segundo sua função de ativação, emitem o sinal resultado à camada posterior, uma à outra até a camada de saída. Os neurônios da camada de saída (*yi*) emitem o sinal resultado do padrão apresentado como entrada.

Os principais elementos utilizados na descrição de RNAs são a representação distribuída, as operações locais e o processamento não-linear. Estes atributos especificam duas aplicações básicas das RNAs: situações onde poucas decisões tem que ser tomadas a partir de uma grande quantidade de dados e situações onde um complexo mapeamento não-linear deve ser aprendido.

A aprendizagem da rede dá-se pela modificação sistemática dos pesos sinápticos das conexões entre os neurônios das diversas camadas, de modo a armazenar o conhecimento apresentado à rede.

Um dos algoritmos de aprendizagem mais utilizados é conhecido como Propagação Retroativa do Erro (*backpropagation error*), sua descrição foi desenvolvida por RUMELHART et al. (1986).

O Back-Propagation é um algoritmo supervisionado que utiliza pares de entradas e saídas desejadas, para ajustar os pesos da rede por um mecanismo de correção erros. O aprendizado ocorre em duas etapas: na primeira os sinais de entrada são apresentados à rede, propagados pelas diversas camadas até a camada de saída, onde é comparado com a saída desejada. O erro gerado pela saída é então retropropagado e os pesos das conexões reajustados. A Equação 1 mostra o cálculo do erro para a apresentação de todos os padrões *p* de entrada.

$$E = \frac{1}{2}\sum_{p}\sum_{i=1}^{k}(Xo_i^p - Xe_i^p)^2 \quad (1)$$

Símbolos	
E	erro total da rede
Xo_i^p	valores observados na saída i
Xe_i^p	valores estimados na saída i
p	número de padrões
k	número de unidades de saída

O ajuste feito em um peso sináptico de um neurônio é proporcional ao produto do sinal de erro pelo sinal de entrada da sinapse em questão. Dessa forma, HAYKIN (2000), definem a correção dos pesos como:

$$w_{kj}(t+1) = w_{kj}(t) + \eta\delta_j(t)x_k(t) \quad (2)$$

Símbolos	
$w_{kj}(t+1)$	peso na conexão kj no tempo t+1
$w_{kj}(t)$	peso na conexão kj no tempo t
η	taxa de aprendizagem da rede
$\delta_j(t)$	erro do neurônio j no tempo t
$x_k(t)$	valor da entrada k no tempo t

A taxa de aprendizagem η (pronuncia-se "eta") da rede neural, como fator multiplicativo na correção dos pesos das conexões entres os neurônios, indica a rapidez como os pesos sinápticos são ajustados, ou em outras palavras, a velocidade de aprendizagem da rede.

À medida que a taxa de aprendizagem aumenta, o tempo necessário para a rede neural adquirir certo volume de conhecimento tende a diminuir. No entanto, um aumento excessivo na taxa de aprendizagem, pode acarretar instabilidade, pois os pesos seriam corrigidos abruptamente prejudicando a convergência da rede, isto é, a minimização do erro.

Como descreve SILVA (2002), o processo de aprendizagem pode ser divido em duas etapas, conhecidas como fase de treinamento e fase de teste. Os padrões utilizados no aprendizado das redes neurais devem ser divididos em dois conjuntos de intersecção vazia, um para cada uma das fases.

O conjunto de treinamento, que engloba geralmente 75% dos padrões, é utilizado para a correção dos pesos das conexões, visando alcançar os valores desejados para a saída.

O conjunto de teste, contendo os 25% restantes, é utilizado apenas para calcular o erro total da rede, o qual objetiva-se minimizar. Por não alterar os pesos, a fase de teste não permite que a rede armazene o conhecimento sobre os padrões de teste, o qual é utilizado tão somente para verificar a quantidade e a qualidade do conhecimento armazenado pela rede, durante o treinamento.

Dessa forma, amplia-se a expectativa da rede neural gerar bons resultados para padrões semelhantes, mas não idênticos aos do conjunto completo, pois uma parte dele foi utilizada para ensinar e outra para avaliar. É comum dividir os padrões de entrada em 3 conjuntos distintos: treinamento (60%), teste (20%) e validação (20%). Medindo o erro dos padrões de validação, que também não alteram os pesos da rede, é possível avaliar melhor o aprendizado da rede.

3. MATERIAIS E MÉTODOS

Foi utilizado para os testes um *software* especialmente desenvolvido para esse artigo com base em *Borland Delphi 7.0* que se utiliza de componente próprio para a construção de redes neurais disponibilizado em Medeiros (2000).

O menu de Geral de entrada do programa, em Parâmetros, tem uma interface que permite a manipulação do número de épocas, a taxa de inércia, a taxa de aprendizagem, o número de neurônios na camada oculta e o arquivo onde se gravam as experiências simuladas na fase de teste e treinamento. É importante notar que esse arquivo representa o conhecimento da rede, ou seja, o que foi aprendido até então.

Em Processamento, encontra-se o núcleo do programa. Primeiro tem-se a opção Conjunto de Entrada. Nessa opção é onde se escolhe a tabela que contém o conjunto de dados que será usada para treinar a rede. Três tabelas estão disponíveis: uma com dados de 2001 apenas; outra com dados de 2001 e 2002; e ainda outra, com dados das turmas de 2001, 2002 e 2003.

A próxima opção desse menu é Treinamento. Aqui se encontram dois botões: Construir e Treinar. Basicamente, ao clicar em Construir o programa constrói uma nova rede com os dados configurados em Parâmetros do menu Geral. E o botão Treinar deve ser acionado para que a rede seja automaticamente treinada.

Por fim, a opção Teste permite que façamos testes com dados das três turmas usadas no treinamento e, ainda, com a turma de 2004. Permite ainda que se escolha a tabela de destino dos resultados da previsão da rede. Nessa tela pode-se fazer tanto testes simples, ou seja, individuais para os alunos ou para dados quaisquer, quanto testes mais complexos, como o teste de uma turma inteira. O teste simples é acionado ao se inserir o valor desejado de cada parâmetro de entrada e clicar o botão Teste Simples. Para o teste mais elaborado, basta escolher a tabela de teste desejada e clicar em Teste Complexo.

Neste programa são inseridos os dados de treinamento: os de entrada e os de saída, conforme é mostrado no Quadro 1. Os dados de entrada contém informações de cada aluno sobre sua idade, experiência militar, estado de origem, estado civil e número de filhos, área de atuação, sexo e colocação e nota no concurso de ingresso. Os dados de saída constituem-se da classificação de cada um desses mesmos alunos no final do curso.

Assim, a rede é treinada a fazer uma associação do padrão de características gerais do indivíduo ao ingressar no curso com sua classificação final no mesmo. Essa é a fase necessária de treinamento da rede, em que ela entra em contato com um conjunto de dados que irão capacita-la a julgar o peso de cada dado dentro do conjunto de dados na colocação final do aluno. Então, a rede estará treinada e pronta para identificar os padrões de novos conjuntos de dados de entrada. Dessa forma, a rede poderá realizar os testes desejados em uma nova amostra de dados.

4. TESTES E RESULTADOS

Neste artigo são usados dados de alunos da Escola de Administração o Exército (EsAEx) para os anos de 2001, 2002, 2003 e 2004.

No primeiro teste, a rede será treinada com os dados da turma de 2001 e testada com os da de 2002, que entram no programa sem os dados da classificação final. Ou seja, estes são inseridos na interface do usuário. Inserir-se-á na rede informações de indivíduos (dados 2002) que, a princípio, não se sabe a classificação e espera-se que aquela faça a previsão da classificação dos mesmos de acordo com o padrão de características que aprendeu anteriormente. E depois será comparado o resultado do programa com o resultado real.A rede irá então aprender com os padrões de 2001 e prever as classificações de 2002.

No segundo teste, a rede será treinada com dados de 2001 e 2002 e testada com os dados de 2003 da mesma forma que anteriormente. Com isso, esperamos ver uma evolução no erro de previsão da rede.

Por último, no terceiro teste, para fazer a previsão das classificações dos alunos de 2004, dispõem-se de informações referentes aos padrões de comportamento de três anos anteriores, que a RNA já terá aprendido. E, além disso, pretende-se estimar qual a margem de erro desta previsão comparando com os erros encontrados nos testes anteriores.

Os dados foram matematizados para inserção no programa da seguinte maneira:

Atributos		Domínio	Categoria
PESSOAIS	Sexo	M, F	1, 2
	Formação	Militar, Civil	1, 2
	Estado Civil	Casado, Solteiro	1, 2
	Filhos	S, N	1, 0
	Idade	18-37	-
	Região de Origem	Sudeste Nordeste Norte Sul Centro Oeste Rio de Janeiro	1 2 3 4 5 6
	Especialidade	Seção de Ensino: SE 1 SE 2 SE 3 SE 4 SE 5	 1 2 3 4 5
CONC	Classificação na área	1 –5 (Por curso) 1-20 (Informática)	-
	Classificação Geral	1 – No. Vagas	-

CURSO	Nota Final	5 – 10	-
SAÍDA	Classificação Final do Curso	1 – 10 11 – 25 26 – 50 50	1 2 3 4

Quadro 1 – Dados Utilizados

Foram realizados 70 testes com os parâmetros da rede, conforme Anexo A. Basicamente, uma rede neural possui os seguintes parâmetros:

Estrutura: Estrutura da rede neural, número de neurônios por camada. A presente rede tem 3 camadas. A estrutura da rede ficou estabelecida com 10 neurônios de entrada e 4 de saída. Existe ainda uma camada intermediária, cujo número de neurônios é configurável.

Taxa de aprendizado da rede: velocidade que a rede irá corrigir os dados. A rede lê várias vezes os mesmos dados de forma a fixar os padrões correspondentes. Cada vez que ela lê um conjunto de dados, ela aprende, define e corrige os pesos de cada dado pertencente a esse conjunto. Uma alta taxa de aprendizagem significa que a rede irá corrigir o peso dado a um parâmetro de uma forma mais sensível. Uma taxa baixa torna essa correção mais demorada.

Taxa de Inércia: A taxa de inércia é um parâmetro que diz o quanto do conhecimento passado ela usará em seu teste presente. É a taxa que a rede conserva o seu conhecimento passado. Uma taxa de inércia maior evita que a rede pare em mínimos locais.

Nº. Épocas*:* Número de vezes que a rede lê os dados para aprende-los. Cada vez que os dados são lidos, a rede corrige seus parâmetros de forma a desenvolver e consolidar um padrão mais coerente para classificar o conjunto de dados.

Erro Total: É o mínimo erro quadrado. A rede calcula essa taxa com base na diferença no valor esperado do dado (calculado pela rede treinada já com o primeiro conjunto de dados) e o valor obtido.

A Tabela 2 mostra o melhor resultado encontrado, ou seja, aquele que gerou o menor Erro Total da rede e cuja configuração será utilizada nos testes:

Parâmetro	Valor
No. Neurônios da camada oculta	2,00
Tx de Aprendizado	0,10
Tx de Inércia	0,60
No. De Épocas	510
Erro Total	0,20

Tabela 2 – Configuração da Rede

5. DISCUSSÃO

Conforme descrito anteriormente, foram realizados três testes e obteve-se o seguinte resultado de Erro Total da rede (utilizou-se a mesma estrutura para todos, descrita na Tabela 2), como é apresentado na Tabela 3:

Teste	Dados Testados	Erro Total
1º.	2002	0,183850
2º	2003	0,208006
3º	2004	0,208808

Tabela 3 – Resultado Erro Total

Observa-se que o Erro manteve-se praticamente igual em todos os testes quando deveria ter diminuído, pois, a rede trabalhou com mais dados de treinamento de um teste para outro. Uma possível explicação é de que o conjunto de padrões provavelmente não varie muito de um ano para outro.

A próxima experiência realizada foi comparar o resultado obtido pela rede com o resultado real para os anos de 2002 e 2003. Em seguida, calculou-se um índice de acertos da rede conforme a expressão 3:

$$IA = \frac{\text{No. de classificações certas}}{\text{No. de classificações totais}} \times 100 \quad (3)$$

Para o ano de 2002, obteve-se um IA de 47,5% e para o ano de 2003, um índice de 41,34%. Mais uma vez, mesmo com mais dados, a rede não conseguiu melhorar o seu desempenho.

Apesar do Erro Total da rede em prever os dados ser baixo (20,00%), ao compararmos o resultado desta com a situação real verificou-se falta de precisão que pode estar relacionada à qualidade dos dados, ou seja, a forma como estes foram trabalhados, ou à eficiência da rede em fazer este tipo de previsão. Deve-se ressaltar que neste artigo estão sendo usados dados apenas de quando o aluno entra no curso e de quando sai. Nenhum dado de meio de curso foi utilizado, portanto não se contabiliza dados de evolução.

Cabe lembrar ainda que outros trabalhos na área de redes neurais conseguiram uma taxa de erro em torno de 20% também. Um exemplo é o já mencionado trabalho de Silva (2002) que consegue atingir um erro de 21%.

Para o ano de 2004, o qual ainda não se têm os dados reais, e calculando uma média dos índices acima, prevê-se que a rede acertará em torno de 79,02% para uma rede treinada com 510 épocas, uma taxa de aprendizagem de 10%, uma taxa de inércia de 60% e uma estrutura 10-4-4, ou seja, 10 neurônios de entrada, 4 intermediários e os 4 de saída.

Pode-se comparar este resultado a um jogo de azar, onde a probabilidade de se ganhar é de 0,25. Pois, tem-se um grupo de alunos que no final de um ano de instrução serão classificados de ordem ordenada em quatro classes de desempenho. Após realizado um teste de hipótese conforme Anexo B, verificou-se que a rede é estatisticamente viável o que torna evidente a possibilidade de uso das redes neurais como forma de instrumento auxiliar na preparação educacional de alunos em escolas militares.

Sempre é importante lembrar que a rede aprende a cada nova entrada de dados que fornecemos ao processo. Dessa forma, enquanto a taxa de erro para os dados de treinamento pode ser bastante reduzida, pois a rede é treinada e testada com os mesmos dados simultaneamente em sua fase de treinamento, quando apresentamos à rede um novo conjunto de dados é possível que a taxa de erro varie sensivelmente. Fazendo uma comparação com uma rede treinada com vários cachorros da mesma raça de acordo com suas características biométricas, por mais que no treinamento o erro quadrático médio tenha sido bem pequeno, quando submetemos a rede a um grupo de cavalos possivelmente a rede apresentará um erro maior de identificação. Para tanto faz-se necessário a disponibilidade de maiores informações para a redução do erro da rede. Possivelmente com um número maior de turmas (anos), o resultado obtido será mais preciso.

Por último fez-se uma breve análise qualitativa das características que mais ou menos influenciam na classificação final do aluno. Essa análise foi feita de uma forma bem simples: manteve-se as características de entrada constante, alterando na interface do programa uma ou outra característica e verificando o efeito dessa mudança.

Este efeito foi categorizado para fins de análise da seguinte forma: no 1º. Grupo encontra-se as características que influenciam no grupo de classificação, por exemplo, o aluno pode passar do grupo D para o grupo A ou B. No 2º. Grupo, as características que interferem apenas nas chances do indivíduo ser classificado no grupo A ou B. Por exemplo, se ele tinha 20% de chance de estar no grupo A, com uma mudança de determinado perfil ele passa a ter 40%. E um 3º. Grupo que contem itens que não influenciam ou tem pouca influencia no resultado.

Deve-se ressaltar que a rede obtém seus resultados de acordo com um conjunto de padrões anteriormente aprendidos e correlacionados.

O meio de origem (militar ou civil), a região de origem, o estado civil e classificação na área estão inseridos dentro no 1º grupo de características. O meio militar predomina sobre o civil principalmente para os homens, para as mulheres não há alteração. As regiões que mais contribuem para uma boa classificação são o Rio de Janeiro, o Sul e o Nordeste. Constatou-se que o Estado Civil para as mulheres tem enorme importância, uma vez que as casadas têm mais chance de obter uma melhor classificação.

Finalmente, observou-se que a classificação na área (concurso) possui uma relação inversa com a classificação final (curso).

Os itens que podem ser classificados no 2º. Grupo é: filhos e sexo. Os homens predominam sobre as mulheres, enquanto que os filhos influenciam positivamente, portanto mais para os homens do que para as mulheres. Assim, reunindo os parâmetros analisados, percebe-se que os homens, militares, casados e com filhos formam um perfil predominante. Muito possivelmente, este resultado deve-se a própria cultura militar que valoriza o aspecto familiar.

No 3º. Grupo encontra-se as características de especialização (seção de ensino), Classificação Geral e Nota de Entrada, os dois últimos referentes ao concurso de admissão. Estes itens exerceram pouquíssima influência no resultado.

6. CONCLUSÃO

Neste trabalho, pretendeu-se desenvolver um novo método, para orientar aos cursos de formação militar das Forças Armadas Brasileira, que pudesse associar o perfil dos alunos com seu desempenho de final de curso ainda no início do mesmo, utilizando um instrumento moderno e de alto padrão tecnológico. O uso de redes neurais ainda pode ser muito bem explorado em processos pedagógicos e já nesse trabalho é notável seu índice de acerto em relação a um procedimento aleatório.

Além do mais, existem muitas outras vantagens na adoção de instrumentos de reconhecimento de padrões em estabelecimentos de ensino. As redes neurais são flexíveis. Como pôde ser visto nesse artigo, pude-se fazer testes e manipular dados quantitativos e qualitativos com uma taxa de acuracidade teórica de até 79%. O programa pôde reconhecer padrões de mais de nove variáveis diferentes de mais de três centenas de alunos testados. Cabe destacar também sua capacidade de lidar com o abstrato. Quando se testam parâmetros como a influência do estado civil ou da nota inicial do concurso na classificação final, fica clara a habilidade da rede em lidar com conceitos abstratos e subjetivos.

Além disso, a rede é ainda tolerante a falhas ou casos isolados já que o treinamento da rede se dá de forma a perceber um padrão que leva em consideração todos os casos disponíveis para treinamento. Dessa forma, por mais que se tenha um caso destoante no treinamento da rede, ainda assim a atribuição de pesos dos parâmetros será feita de tal forma que o impacto do caso especial seja mínimo. E quanto maior o número de dados disponíveis para testes, menor o impacto do caso diferente. Doravante, podemos ainda definir uma nova classe de parâmetro (dados de saída) para o caso especial para aumentar o índice de acerto da rede.

Conclui-se que vários fatores poderiam ser mais bem trabalhados, como, por exemplo, mais testes de configuração da rede, de dados de entrada e saída e uma análise estatística mais profunda nos perfis predominantes e nos características que influenciam nos resultados obtidos.

Muito importante seria obter bancos de dados mais amplos e com maior qualidade, por parte das instituições de ensino, de atributos que poderiam alterar ou complementar os resultados alcançados pela rede.

Dentro das condições de pesquisa o resultado foi bastante satisfatório. Com um índice de acerto teórico de mais de 75% e um índice de acerto real de mais de 45% (quando o esperado no teste aleatório é de 25%), torna-se evidente a funcionalidade da rede desenvolvida pelo grupo. E mais, verificou-se a importância de conhecer o aluno antes mesmo deste chegar no Curso. Além disso, destaca-se a essencialidade de acompanhar a evolução do respectivo indivíduo durante seu processo de aprendizagem e aprimoramento.

Os obstáculos existentes para o uso das redes neurais, verificados no contexto do artigo, estão mais associados à quantidade de dados disponíveis para treinamento e teste do que a questões qualitativas como processamento de dados, custo, desenvolvimento teórico ou disponibilidade de máquina. Claro, outra questão central no uso e adoção de redes neurais na previsão de desempenho está associada à definição dos parâmetros que serão analisados pela rede e na definição das respostas esperadas. Mas isso apenas comprova a flexibilidade e maleabilidade da teoria de redes neurais.

Sugere-se que mais pesquisas sejam feitas nestas áreas, tanto no desenvolvimento de Redes e outros métodos computacionais auxiliando a Educação, quanto na aplicação destes métodos associados a analises psicopedagógicas que possam fornecer subsídios aos orientadores e instrutores de Cursos de Formação Militar e assim estes possam aprimorar o nível de qualidade e eficiência do capital humano das Forças Armadas Brasileira.

7. REFERÊNCIAS BIBLIOGRÁFICAS

[1] RUMELHART, D. E., HILTON, G. E., WILLIAMS, R. J. **Learning Representations by Back-Propagation Error.** Nature, 1986.

[2] HAYKIN, S., **Redes Neurais - Princípios e Prática.** Tradução: Paulo Martins Engel, Editora Bookman, Porto Alegre RS, 2ª Edição, 2000. 900 p.

[3] SILVA, A. F.. **Previsão Da Evapotranspiração de Referência Utilizando Redes Neurais** . 2002, 68p. Dissertação (Mestrado em Engenharia Agrícola) - Curso de Pós-Graduação em Engenharia Agrícola, Universidade Federal de Viçosa, Viçosa.

[4] MEDEIROS, L. F., **Redes Neurais em Delphi.** Editora Visual Books, Florianópolis SC, 1ª Edição, 2003**.** 115p.

Anexo A – Teste dos Parâmetros da Rede:

Teste	Épocas	Camadas Ocultas	Tx. Aprendizagem	Tx. Inercia	Erro
1	510	1	50	10	0,344474
2	510	2	50	10	0,286316
3	510	3	50	10	0,259047
4	510	4	50	10	0,246608
5	510	5	50	10	0,234259
6	510	10	50	10	0,22517
7	100	4	50	10	0,22379
8	200	4	50	10	0,221299
9	300	4	50	10	0,218048
10	400	4	50	10	0,214435
11	500	4	50	10	0,211112
12	510	4	50	10	0,209275
13	700	4	50	10	0,21056
14	510	4	20	10	0,209002
15	510	4	10	10	0,20794
16	510	4	70	10	0,238224
17	510	4	100	10	0,261712
18	510	4	0	10	0,245269
19	510	4	10	5	0,248512
20	510	4	10	20	0,246061
21	510	4	10	15	0,241315
22	510	4	10	10	0,236955
23	510	4	10	30	0,23288
24	510	4	10	40	0,228975
25	510	4	10	50	0,225531
26	510	4	10	60	0,222775
27	510	4	10	80	0,226486
28	510	1	10	60	0,219399
29	510	2	10	60	0,215762
30	510	3	10	60	0,212843
31	510	4	10	60	0,210369
32	100	4	10	60	0,209919
33	200	4	10	60	0,209055
34	300	4	10	60	0,207833
35	400	4	10	60	0,206322
36	510	4	10	60	0,204561
37	510	4	0	60	0,202072
38	510	4	10	60	0,202953
39	510	4	30	60	0,218066
40	510	4	50	60	0,222184
41	510	4	70	60	0,269365
42	510	4	100	60	0,304165
43	510	4	10	10	0,26718
44	510	4	10	30	0,261115
45	510	4	10	60	0,251686

46	510	4	10	60	0,242975
47	510	1	50	10	0,236589
48	400	2	20	70	0,234316
49	400	2	70	20	0,23461
50	500	2	10	60	0,226191
51	500	4	10	60	0,222265
52	1000	4	10	60	0,21553
53	700	4	10	60	0,2122
54	700	4	60	10	0,216704
55	500	2	20	80	0,2266
56	500	2	80	20	0,227756
57	500	2	70	10	0,22566
58	500	2	10	70	0,219948
59	550	2	10	60	0,216029
60	550	2	60	10	0,218899
61	510	2	10	10	0,212139
62	510	2	50	10	0,212787
63	510	1	50	10	0,215429
64	510	2	10	40	0,210715
65	510	2	10	60	0,211012
66	510	2	10	60	0,20888
67	510	2	10	60	0,208006
68	510	2	10	60	0,207202

Anexo A – Teste de Hipótese para a probabilidade de acerto na classificação dos alunos:

Sejam x_1, ..., x_n um conjunto de observações independentes tais que $x_i = \begin{cases} 1, \textit{com probabilidade } p \\ 0, \textit{com probabilidade } 1-p \end{cases}$, com 0 < p < 1 e i = 1, ..., n. Então, temos que a variável y = x_1 + ... + x_n segue uma distribuição binomial (n, p), onde n é o número de observações e p a probabilidade de sucesso. Assim, podemos definir a proporção de sucessos na amostra como:

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

Usando o Teorema do Limite Central, temos que:

$$\overline{X} \overset{a}{\sim} N(p, \frac{p\,(1-p)}{n}),$$

onde $\overline{X}$ tem distribuição aproximadamente normal com média p e variância $\frac{p\,(1-p)}{n}$.

Como não conhecemos o valor de p, queremos testar se a proporção de sucessos na amostra é ou não estatisticamente diferente de um dado valor p_o:

H_o: p = p_o
H_1: p > p_o

O teste é construído pela análise da probabilidade de se cometer erro Tipo I (rejeitar H_o quando H_o é verdadeira). Se esta probabilidade é baixa, isto significa que podemos rejeitar H_o pois a chance de cometer erro é pequena. Usualmente, H_o é rejeitada para valores de p menores que 5%.

Cálculo de p:

Dada uma amostra, calculamos $\bar{x}$ (valor observado da variável $\overline{X}$). A questão é: se H_o é verdade, qual a probabilidade de observarmos valores iguais ou maiores que $\overline{X}$ (ou seja, de rejeitarmos H_o a partir do valor amostral observado)?

$$P = P_r[\text{Re j. } H_o \,|\, H_o \text{ verdade}] = P_r[\overline{X} > \bar{x}]$$

$$P_r[\overline{X} > \bar{x}] = P_r\left[\frac{\overline{X} - p_o}{\sqrt{\frac{p_o(1-p_o)}{n}}} > \frac{\bar{x} - p_o}{\sqrt{\frac{p_o(1-p_o)}{n}}}\right] = P_r\left[Z > \frac{\bar{x} - p_o}{\sqrt{\frac{p_o(1-p_o)}{n}}}\right],$$

sendo Z uma variável com distribuição normal de média 0 e variância 1 (normal padrão).

Aplicação:

A probabilidade p é a proporção de acertos da rede neural. Se tentarmos adivinhar aleatoriamente em qual dos quatro grupos de desempenho um componente de uma turma de 100 pessoas está, teremos que:

H_o: p = $\frac{1}{4}$

H_1: $p > \frac{1}{4}$

Assim temos:

$\overline{X} \overset{a}{\sim} N(\frac{1}{4}, \frac{\frac{1}{4}(1-\frac{1}{4})}{100}) \overset{a}{\sim} N(\frac{1}{4}, \frac{3}{1600})$.

Então, testamos $\overline{x}$ = 0,79 (se um índice de acerto teórico de 79% é estatisticamente provável):

$$P = P_r[Z > \frac{0{,}79 - \frac{1}{4}}{\sqrt{\frac{3}{1600}}}] = P_r[Z > 12{,}47] = 0{,}0000$$

E $\overline{x}$ = 0,45 (se um índice de acerto real de 45% é estatisticamente provável):

:

$$P = P_r[Z > \frac{0{,}45 - \frac{1}{4}}{\sqrt{\frac{3}{1600}}}] = P_r[Z > 4{,}62] = 0{,}0000$$

Em ambos os casos rejeitamos H_o. Isto é, a probabilidade de acerto da rede é maior que o acaso.