



Estudo Comparativo de Técnicas de Redes Neurais Artificiais na Previsão da Velocidade do Vento em Curto Prazo

Trabalho de Conclusão de Curso

Engenharia da Computação

Carlos Vinicius Aladim Veras
Orientador: Mêuser Jorge Silva Valença



UNIVERSIDADE
DE PERNAMBUCO

**Universidade de Pernambuco
Escola Politécnica de Pernambuco
Graduação em Engenharia de Computação**

Carlos Vinicius Aladim Veras

**Estudo Comparativo de Técnicas de
Redes Neurais Artificiais na Previsão
da Velocidade do Vento em Curto
Prazo**

Monografia apresentada como requisito parcial para obtenção do diploma de Bacharel em Engenharia de Computação pela Escola Politécnica de Pernambuco – Universidade de Pernambuco.

Recife, Novembro de 2013.

MONOGRAFIA DE FINAL DE CURSO

Avaliação Final (para o presidente da banca)*

No dia 17 de 12 de 2013, às 11:00 horas, reuniu-se para deliberar a defesa da monografia de conclusão de curso do discente CARLOS VINICIUS ALADIM VERAS, orientado pelo professor Meuser Jorge Silva Valença, sob título Estudo comparativo de técnicas de redes neurais artificiais na previsão da velocidade do vento em curto prazo, a banca composta pelos professores:

Sérgio Murilo Maciel Fernandes

Meuser Jorge Silva Valença

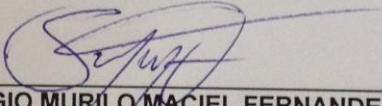
Após a apresentação da monografia e discussão entre os membros da Banca, a mesma foi considerada:

Aprovada Aprovada com Restrições* Reprovada

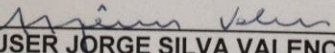
e foi-lhe atribuída nota: 9,5 (NOVE E MEIO)

*(Obrigatório o preenchimento do campo abaixo com comentários para o autor)

O discente terá 03 dias para entrega da versão final da monografia a contar da data deste documento.



SÉRGIO MURILO MACIEL FERNANDES



MEUSER JORGE SILVA VALENÇA

* Este documento deverá ser encadernado juntamente com a monografia em versão final.

*Dedico este trabalho aos meus pais, que me incentivaram e me apoiaram durante a
minha vida.*

Agradecimentos

Agradeço aos meus pais, Carlos e Eldy Veras, pelo apoio, confiança e educação.

Agradeço aos amigos que participaram ativamente de todo meu processo de graduação, sempre me aconselhando e me ajudando, em especial a Bruna Galle pelo grande apoio na realização deste trabalho.

Finalmente agradeço ao meu orientador, Mêuser Valença, primeiramente por ser o introdutor da área de redes neurais no meu percurso acadêmico, área que muito me identifiquei, e também pela orientação, apoio e paciência.

Resumo

O recurso eólico apresenta grande potencial como alternativa energética. O crescimento de sua capacidade instalada impulsiona a busca por melhorias na previsão da velocidade do vento, que possui caráter variável e dinâmico. Devido a esse caráter o maior desafio é prever a velocidade em um instante futuro. Essa previsão é importante para determinar o potencial eólico futuro, ajudando o operador elétrico a despachar a energia de maneira otimizada além de auxiliar na organização da agenda de manutenção. Uma das formas mais comuns de previsão com métodos computacionais é a utilização de algoritmos de computação inteligente, conhecidas como Redes Neurais Artificiais, e outra forma bastante comum para previsão é utilizar o modelo de persistência, que apresenta bons resultados para previsões de muito curto prazo. Este trabalho se propõe a usar diferentes configurações de três tipos de redes neurais, o Multilayer Perceptron, as redes recorrentes de Elman e as redes recorrentes de Jordan, para prever a velocidade do vento. Além disso, serão realizados estudos comparativos dos resultados de cada rede com modelos de referência utilizados atualmente (modelo de persistência), com a finalidade de apontar possibilidade de melhorias e diminuições de erro de previsão.

Palavras-chave: Previsão de vento, Redes Neurais Artificiais, Multilayer Perceptron, Redes Neurais Recorrentes de Elman, Redes Neurais Recorrentes de Jordan

Abstract

The wind resource has great potential as an energy alternative. The growth of its installed capacity drives the search for improvements in forecasting the wind speed, which has variable and dynamic character. Due to this character's greatest challenge is to predict the speed at a future time. This prediction is important in determining the future wind potential, helping the operator to dispatch electrical energy optimally. One of the most common forms of computational prediction methods is the use of intelligent computation algorithms, known as Artificial Neural Networks, and another common way to use the model forecast is persistence, which presents good results. This paper proposes to use different configurations of three types of neural networks, Multilayer Perception, Elman recurrent neural networks and Jordan recurrent neural networks, to predict wind speed. Furthermore, comparison of the results of each network reference models currently (persistence model), used for the purpose of pointing out the possibility of improvements and reductions in forecast error studies will be conducted.

Keywords: Wind Prediction, Artificial Neural Network, Multilayer Perceptron, Elman Recurrent Neural Network, Jordan Recurrent Neural Network

Sumário

Capítulo 1 - Introdução	1
1.1 Motivação	1
1.1.1 Objetivos.....	2
1.2 Estrutura da Monografia	3
Capítulo 2 - Fundamentação Teórica	4
2.1 Energia Eólica	4
2.1.1 História	4
2.1.2 Conceitos.....	6
2.1.3 Turbina Eólica.....	7
2.1.4 Potencial Eólico Brasileiro	9
2.1.5 Importância dos modelos de previsão	10
2.1.6 Modelo de Persistência	11
2.2 Redes Neurais Artificiais	12
2.2.1 Neurônio Biológico.....	12
2.2.2 Neurônio Artificial	13
2.3 Multilayer Perceptron.....	14
2.3.1 Algoritmo Backpropagation.....	16
2.3.2 Critério de Parada do Treinamento da Rede	18
2.4 Redes Neurais Recorrentes de Elman	19
2.4.1 Topologia da Rede	20

2.4.2	Treinamento da Rede de Elman	21
2.5	Redes Neurais Recorrentes de Jordan	22
2.6	Encog	23
Capítulo 3 - Metodologia		24
3.1	Base de Dados	24
3.2	Pré-Processamento dos Dados	24
3.3	Normalização dos Dados	25
3.4	Conjuntos de Dados	25
3.5	Desnormalização dos Dados	26
3.6	Implementação das RNAs	27
3.6.1	Camada de Entrada	27
3.6.2	Camada Escondida	27
3.6.3	Camada de Saída	28
3.6.4	Camada de Contexto	28
3.6.5	Função de ativação	28
3.6.6	Taxa de Aprendizado	28
3.6.7	Momentum	29
3.6.8	Critério de Parada	29
3.6.9	Treinamento da RNA	29
3.7	Teste Comparativo	29
3.7.1	Teste de Igualdade	30

3.7.2	Comparação entre as RNAs.....	30
3.7.3	Comparação com o Modelo de Persistência	31
Capítulo 4 - Resultados		32
4.1	Estudo de Sensibilidade da Camada de Entrada	32
4.2	Escolha da Quantidade de Neurônios na Camada Escondida	37
4.3	Testes Comparativos com o Modelo de Persistência.....	38
Capítulo 5 - Considerações Finais		40
5.1	Conclusões.....	40
5.2	Trabalhos Futuros	41
Bibliografia		42

Índice de Figuras

Figura 1. Aerogeradores de Eixo Horizontal.....	7
Figura 2. Aerogerador de Eixo Vertical.....	8
Figura 3. Neurônio Biológico	13
Figura 4. Modelo Matemático de um Neurônio Biológico	14
Figura 5. Arquitetura da Rede MLP	15
Figura 6. Conjuntos linearmente separável e não linearmente separável	15
Figura 7. Gráfico do melhor ponto de parada do treinamento	19
Figura 8. Topologia da Rede de Elman	21
Figura 9. Topologia da Rede de Jordan	22
Figura 10. Simulações da camada escondida.....	38

Índice de Tabelas

Tabela 4.1 Erros Médios Quadráticos para a rede MLP	33
Tabela 4.2 Erros Médios Quadráticos para da rede de Elman	34
Tabela 4.3 Erros Médios Quadráticos para a rede de Jordan	35
Tabela 4.4 Média dos erros médios quadráticos de todas as RNAs.....	36
Tabela 4.5 Comparação das entradas nas redes MLP	36
Tabela 4.6 Comparação das entradas nas redes de Elman	37
Tabela 4.7 Comparação das entradas nas redes de Jordan	37
Tabela 4.8 Comparação entre as arquiteturas das RNAs.....	39
Tabela 4.9 Comparação das RNAs com o modelo de persistência	39

Tabela de Símbolos e Siglas

EMQ – Erro Médio Quadrático

MLP – Multilayer Perceptron

RNA - Rede Neural Artificial

Capítulo 1

Introdução

Este capítulo inicia com a motivação para a realização deste Trabalho de Conclusão de Curso (TCC). Em seguida, são definidos os objetivos a serem alcançados. Por fim, detalha os conteúdos a serem abordados nos capítulos seguintes.

1.1 Motivação

Um dos grandes problemas do mundo de hoje é a questão relativa à energia. O consumo de energia é uma variável crescente em escala mundial. Portanto, o desafio energético mundial é satisfazer esta crescente demanda de energia, proteger as fontes e, ao mesmo tempo, afrontar a ameaça igualmente urgente da mudança climática.

O aproveitamento de energia ainda não atingiu um nível satisfatório, visto que a imensa maioria de energia utilizada no planeta é de origem não renovável, seja de fonte mineral, atômica ou térmica. A energia pode ser utilizada de forma mais civilizada e menos dispendiosa. Uma dessas alternativas, a energia eólica, que causa menor impacto ambiental, é segura e disponível em quase todos os países do mundo. Também há outras fontes renováveis como a energia solar, das marés, geotérmica, das águas e outras mais. As energias renováveis prometem ser fundamentais para a humanidade. Elas são seguras, limpas e abundantes. Ao contrário dos combustíveis convencionais, as energias renováveis têm disponibilidade em quase todos os países do mundo, sendo a energia eólica a de maior interesse na atualidade. [1]

O desenvolvimento da humanidade através da utilização cada vez maior de fontes renováveis de energia minimizará os impactos ambientais. Em particular,

melhorar a eficiência na produção de eletricidade significa melhorar o aproveitamento da energia disponível nos ventos para a produção de eletricidade e isto é de suma importância para a maioria dos países desenvolvidos e em desenvolvimento.

Embora o mercado de usinas eólicas ainda esteja vivenciando a primeira onda de crescimento no Brasil, ele já movimentava 2 bilhões de dólares no mundo. Considerando o grande potencial eólico do Brasil, confirmado através de estudos recentes, é possível produzir eletricidade a custos competitivos com centrais termoelétricas, nucleares e hidroelétricas [2].

A previsão da velocidade do vento é importante para ajudar no planejamento e na melhor utilização dos recursos disponíveis, além de diminuir a incerteza e inconstância na geração da energia. Portanto o maior desafio nessa previsão é encontrar um modelo que melhor consiga prever a velocidade do vento, levando em consideração as suas características altamente voláteis.

1.1.1 Objetivos

O objetivo principal do trabalho proposto é fazer um estudo comparativo entre três configurações, utilizando duas técnicas distintas de Redes Neurais Artificiais, com o intuito de prever a velocidade do vento em curto prazo.

Os objetivos secundários são:

- Realizar um estudo de sensibilidade das entradas nas três configurações;
- Realizar uma comparação entre os resultados obtidos e os modelos de persistência.

1.2 Estrutura da Monografia

A monografia segue a seguinte estrutura: o Capítulo 2 fornece a fundamentação teórica necessária para se compreender a energia eólica, a importância da previsão na geração desta forma de energia e as Redes Neurais Artificiais utilizadas no trabalho. No Capítulo 3 é demonstrado como os experimentos serão conduzidos, como as Redes Neurais serão configuradas e quais medições estatísticas serão utilizadas. Em seguida, no Capítulo 4, são exibidos todos os resultados obtidos assim como as comparações realizadas. Por fim, no Capítulo 5, serão demonstradas quais dificuldades foram encontradas, discussões geradas e as conclusões obtidas, além de sugestões de possíveis melhorias e trabalhos futuros.

Capítulo 2

Fundamentação Teórica

Este capítulo fornece explicação sobre o conteúdo que será abordado no trabalho. A seção 2.1 aborda a energia eólica. A seção 2.2 discorre sobre Redes Neurais Artificiais (RNAs). A seção 2.3 faz considerações a uma arquitetura específica de uma RNA, o Multi-Layer Perceptron (MLP). A seção 2.4 aborda as Redes Recorrentes de Elman, também uma arquitetura específica de uma RNA. A seção 2.5 fala sobre as Redes Recorrentes de Jordan, a última das três arquiteturas utilizadas. A seção 2.6 fala sobre o Encog [20] [21], ferramenta utilizada para implementação das RNAs.

2.1 Energia Eólica

2.1.1 História

A utilização de energia eólica remonta de remotos períodos da humanidade. Registros históricos apontam que na Pérsia, em 200 a.C., utilizavam-se cata-ventos para a moagem de grãos ou a irrigação. Acredita-se que por volta de 2000 a.C., na China, os ventos eram usados para a mesma função, porém num sistema mais rudimentar [3]. Os moinhos eram feitos de madeira montadas em uma base de pedra. Eles eram utilizados para dar suporte à agricultura como forma de bombeamento de água, para irrigação e moagem de grãos. O conceito foi difundido por todo o Oriente Médio até chegar à Europa na época das Cruzadas, nos anos 1000 d.C..

. Durante a idade média, os moinhos tiveram papel significativo na Europa, por sua importância na economia agrícola. Na Holanda, a energia eólica era utilizada dos moinhos na drenagem de água. Durante os séculos XVII ao XIX, cerca de 9000 moinhos existiam em pleno funcionamento na Holanda. O moinho de vento foi aproveitado e sua estrutura foi aperfeiçoada para o modelo tradicional de quatro pás,

que foi usado para a produção de óleos vegetais, papel e para o suporte às serrarias, em 1586. A utilização dos moinhos foi tão grande que em meados do século XIX, foram catalogados também, cerca de 3.000 na Bélgica, 10.000 na Inglaterra e cerca de seiscentos e cinquenta na França [1]. Outro marco importante para a energia eólica na Europa foi a Revolução Industrial. Após a ampliação do uso de máquinas a vapor, observa-se a redução de moinhos em funcionamento na Europa, mas isso não significou o fim de seu uso.

O primeiro dispositivo voltado para transformação da energia eólica em energia elétrica foi desenvolvido nos Estados Unidos. Esse aerogerador foi construído para carregamento de baterias e fornecimento de energia para lâmpadas [1].

A geração de eletricidade através do vento começou em torno do início do século XX e alguns dos primeiros desenvolvimentos foram feitos pelos dinamarqueses. Em 1930, algumas empresas americanas estavam produzindo e comercializando primitivos geradores eólicos. Estas máquinas poderiam fornecer até mil watts de corrente contínua enquanto o vento estivesse soprando [6].

Muitos países europeus começaram a construir desenhos avançados de geradores eólicos, durante os anos 1950 e 1960. A primeira turbina eólica comercial ligada à rede elétrica pública foi instalada em 1976, na Dinamarca. Uma das mais memoráveis máquinas de vento foi construída perto de Rutland, Vermon, nos Estados Unidos e foi projetada para fornecer mil duzentos e cinquenta quilowatts para a malha elétrica de Vermont [6].

No Brasil, a primeira turbina de energia eólica foi instalada em Fernando de Noronha em 1992. Dez anos depois, parques eólicos foram instalados no Rio Grande do Sul (2002), Bahia (2002) e Rio de Janeiro (2003), quando o governo criou o Programa de Incentivo às Fontes Alternativas de Energia Elétrica (Proinfa). Ao total são 59 parques eólicos atualmente em operação em todo o país. Contudo, segundo o Balanço Energético Nacional de 2012 estudos e mapeamentos realizados em todo o território nacional apontam grande potencial eólico ainda não explorado [4].

2.1.2 Conceitos

A energia eólica provém da energia solar. Isto é explicado a partir do fato de que a Terra é aquecida de maneira desigual pelo Sol. Durante o dia o ar sobre a terra se aquece muito mais rápido do que o ar que está sobre as águas. À medida que este ar sobre a terra sobe, o ar mais frio (resfriado pela água) se movimenta para ocupar o espaço deixado por este ar quente. Ao movimento da massa de ar, a partir desta diferença de pressão atmosférica e do movimento do próprio planeta, dar-se-á o nome de vento. A energia eólica é, então, aquela obtida pelo movimento do ar (vento) e é uma fonte abundante, renovável, limpa e disponível em todos os lugares [1].

Como as pressões atmosféricas estão diretamente relacionadas à radiação solar é plausível considerar o vento e, conseqüentemente, a energia eólica uma modalidade de energia solar.

A velocidade instantânea do vento V é calculada como o valor médio acrescido de um desvio a partir da média (flutuação), como pode ser visto na equação 2.1:

$$V(t) = \Delta V + V'(t) \quad (2.1)$$

Onde ΔV corresponde à variação da velocidade média do vento e $V'(t)$ é a flutuação no período t .

Além da turbulência há outros fatores que influenciam na velocidade do vento, como por exemplo, a rugosidade, que consiste na fricção do vento com a superfície terrestre. Quanto maior a rugosidade maior o abrandamento do vento. Regiões florestais ou urbanas abrandam muito o vento, possuem, portanto, alta rugosidade, enquanto que a superfície do mar ou de um lago tem uma rugosidade quase nula. Os obstáculos diminuem a velocidade do vento de forma significativa e normalmente

criam turbulências ao redor deles. Condições climáticas, densidade do ar e relevo também podem causar efeito de aceleração ou desaceleração no escoamento do ar.

2.1.3 Turbina Eólica

Turbinas eólicas, ou aerogeradores, são aquelas que utilizam a energia cinética do vento e a converte em energia mecânica ou energia elétrica. Isso ocorre no momento em que o vento move-se entre as pás da turbina, transferindo parte de sua energia e rotacionando-as.

As turbinas podem ser classificadas quanto à posição do eixo de rotor como turbinas eólicas de eixo horizontal, conforme a Figura 1, onde as pás giram em um plano perpendicular à direção do vento e turbinas eólicas de eixo vertical, conforme a Figura 2, onde as pás giram em um plano paralelo à direção do vento.



Figura 1. Aerogeradores de Eixo Horizontal



Figura 2. Aerogerador de Eixo Vertical

A extração de energia cinética pelas pás do rotor é dada pela energia cinética bruta por unidade de tempo (potência) do vento passando por um rotor de área A , perpendicular ao seu vetor velocidade instantânea V , a qual é dada pela equação 2.2.

$$P = C_p \cdot \frac{1}{2} \cdot \rho \cdot A \cdot V^3 \quad (2.2)$$

Onde:

- C_p = é o coeficiente da potência. É a potência extraída pelo rotor dividida pela potência do vento;
- ρ = densidade do ar;
- A = área do rotor;
- V = velocidade do vento em m/s.

Como se pode ver, a energia potencial depende do cubo da velocidade do vento. Essa forte conexão entre a energia potencial com a velocidade demonstra a importância da previsão do vento para a estimativa precisa da energia disponível.

Diversos fatores contribuem para a complexidade do campo de escoamento das turbinas eólicas, com características não estacionárias, ou seja, características que podem variar ao longo do tempo [6]. Entre eles, os mais importantes são os seguintes:

- Efeitos da camada limite atmosférica modificando a magnitude e a direção da velocidade de corrente livre;
- Giro num plano inclinado em relação ao vento principal, produzindo variações de ângulo de ataque;
- Perturbações no fluxo de suporte e eixo do rotor;
- Deformação estrutural das pás originando velocidades induzidas de translação e rotação dos elementos da pá.

Outro parâmetro importante que deve ser avaliado é a direção do vento, visto que mudanças de direção frequentes indicam rajadas de vento. Além disso, a informação sobre a direção do vento auxilia na determinação da localização das turbinas em um parque eólico [6].

2.1.4 Potencial Eólico Brasileiro

Embora ainda haja divergências entre especialistas e instituições na estimativa do potencial eólico brasileiro, vários estudos indicam valores extremamente consideráveis. Até poucos anos, as estimativas eram da ordem de 20.000 MW. Hoje a maioria dos estudos indica valores maiores que 60.000 MW. A razão dessas divergências decorre principalmente da falta de informações (dados de superfície) e às diferentes metodologias empregadas [24].

De qualquer forma, os diversos levantamentos e estudos realizados e em andamento (locais, regionais e nacionais) têm dado suporte e motivado a exploração comercial da energia eólica no país. Os primeiros estudos foram feitos na Região Nordeste, principalmente no Ceará e em Pernambuco. Com o apoio da ANEEL e do Ministério de Ciência e Tecnologia – MCT, o Centro Brasileiro de Energia Eólica – CBEE, da Universidade Federal de Pernambuco – UFPE, publicou em 1998 a

primeira versão do Atlas Eólico da Região Nordeste. Com o auxílio de modelos atmosféricos e simulações computacionais, foram feitas estimativas para todo o país, dando origem a uma versão preliminar do Atlas Eólico Brasileiro (Figura 5.1), que foi ser concluído pelo CBEE em abril de 2002 [24].

Segundo esses resultados, os melhores potenciais estão no litoral das regiões Norte e Nordeste, onde a velocidade média do vento, a 50 m do solo, é superior a 8 m/s. Entre outras regiões com grande potencial eólico, destacam-se o Vale São Francisco, o Sudoeste do Paraná e o Litoral Sul do Rio Grande do Sul [24].

Também o Centro de Referência para Energia Solar e Eólica – CRESESB/CEPEL elaborou recentemente um atlas eólico. Segundo esses dados, o potencial eólico brasileiro é da ordem de 140.000 MW [24].

2.1.5 Importância dos modelos de previsão

O aumento da participação da energia eólica no cenário mundial cria demandas para o planejamento, operação e mercado de energia. Por conta disso a previsão da geração da energia eólica precisa ser confiável e segura. Como a geração é inconstante é necessário realizar a previsão para garantir ao operador as ferramentas para manter a capacidade de reserva na programação da distribuição de energia de forma a garantir o equilíbrio entre a carga e a geração da mesma, criando reserva de energia para compensar eventuais períodos de geração reduzida.

As previsões podem ser classificadas de acordo com a escala de tempo [6]:

- Curtíssimo prazo: São as previsões de poucos minutos à frente utilizadas principalmente para controlar a potência ativa do gerador;
- Curto prazo: De 1 a 72 horas. Utilizadas para planejar a operação e a distribuição da energia;
- Médio prazo: De 5 a 7 dias. Utilizadas na manutenção das plantas eólicas;

- Longo prazo: Acima de sete dias. Utilizadas em planejamento de recursos energéticos e em leilões de energia

Os modelos de previsão mais utilizados classificam-se em modelos físicos, modelos estatísticos e de inteligência artificial. Os modelos físicos realizam as previsões baseados nos dados meteorológicos e utilizam métodos numéricos para previsão, demandando grande esforço computacional e requerendo um número elevado de dados, sendo recomendados para previsões de médio e longo prazo. Os modelos estatísticos e de inteligência artificial são mais simples e mais rápidos na realização das previsões, contudo possuem tendência à diminuição na precisão da previsão conforme o aumento da escala de tempo. Por essa razão são mais recomendados para previsões de curtíssimo e curto prazo [1]. Dentro dos modelos estatísticos e de inteligência artificial estão as RNAs, que são utilizadas para construir estes modelos.

2.1.6 Modelo de Persistência

Outro modelo que apresenta bons resultados em previsões de velocidade do vento de curto prazo é o chamado modelo de persistência. O modelo de persistência é considerado bastante eficaz na previsão de curta duração de velocidade do vento [15], e tem sido usado como referência para comparação de modelos de previsão. O modelo de persistência é definido como:

$$y(t) = y(t - i) \quad (2.2)$$

Onde $y(t)$ é o valor previsto do vento em um instante t e o termo $y(t - i)$ é o valor da velocidade do vento no instante $t - i$. O modelo de persistência em geral apresenta bons resultados, pois o tempo necessário para que mudanças atmosféricas ocorram é da ordem de várias horas. Para previsão de velocidades de vento em até 6 horas, o modelo de persistência é considerado como a referência a ser superada [16]. Algumas publicações comparam modelos de Redes Neurais

Artificiais (RNA) com o modelo de Persistência, e foi encontrado que normalmente as RNA se sobressaem em relação ao modelo de Persistência, mas que algumas vezes essa vantagem não é muito significativa [17] [18].

2.2 Redes Neurais Artificiais

Redes Neurais Artificiais são técnicas de inteligência artificial inspiradas no funcionamento dos neurônios no cérebro humano. É um sistema adaptativo composto por unidades de processamento interconectadas, chamadas de neurônios, distribuídas em diferentes camadas trabalhando em união para a resolução de um problema. A grande vantagem no uso de redes neurais artificiais para solução de problemas complexos provém, principalmente, de sua capacidade de aprendizagem através de generalização da resposta adquirida durante o treinamento da rede, através do uso de exemplos já ocorridos no passado.

2.2.1 Neurônio Biológico

O neurônio, conforme a Figura 3, é composto pelo corpo celular, dendritos (ramificações do corpo celular), e o axônio (prolongamento do corpo celular), cuja função é transmitir o sinal do corpo celular para as extremidades. As extremidades do axônio são conectadas a outros dendritos pelas sinapses, formando as redes de neurônios. As informações são recebidas pelos dendritos, processadas no corpo celular e enviadas pelos axônios aos outros neurônios. Os neurônios biológicos seguem a Lei do Tudo ou Nada. Isso significa que o estímulo nervoso só é transmitido se ultrapassar o limiar excitatório. Se o estímulo for inferior a esse limiar não ocorrerá impulso nervoso [12].

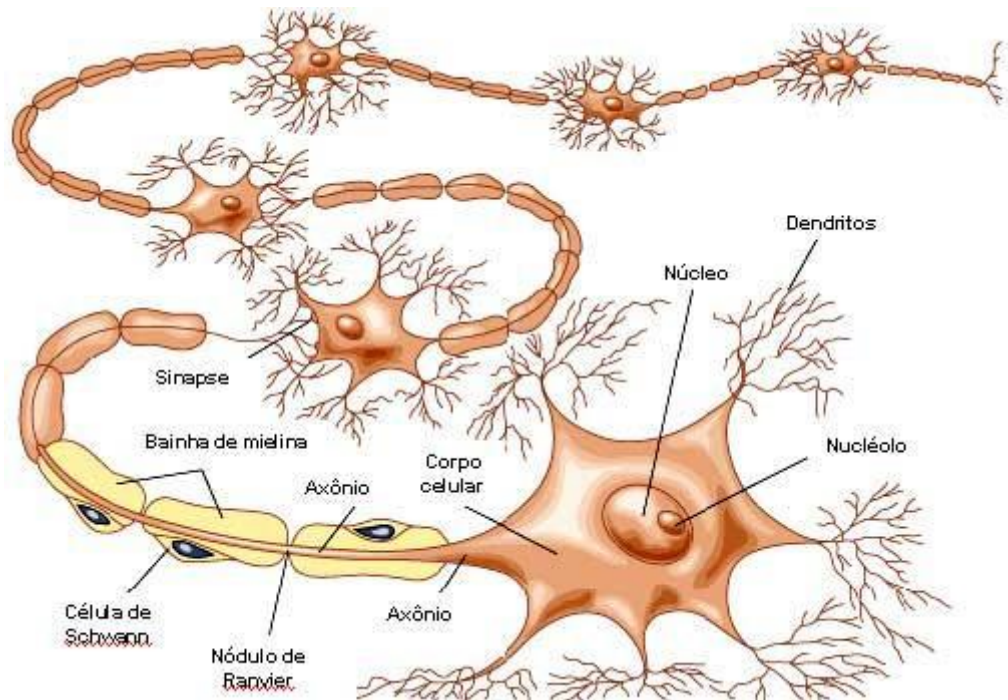


Figura 3. Neurônio Biológico

2.2.2 Neurônio Artificial

O neurônio artificial mais simples, proposto por Mc-Culloch Pitts em 1943, é composto de entradas (que correspondem aos dendritos), unidade de processamento (que corresponde ao corpo celular) e saídas (que correspondem aos terminais do axônio) [12].

A Figura 4 representa matematicamente o neurônio, onde os valores de entrada $x_1, x_2, x_3, \dots, x_n$ são multiplicados pelos pesos sinápticos $w_1, w_2, w_3, \dots, w_n$. O somatório do produto das entradas pelos pesos, dada pela equação 2.4, é passada para a função de ativação $\varphi(\cdot)$, retornando a saída y , dada pela equação 2.5.

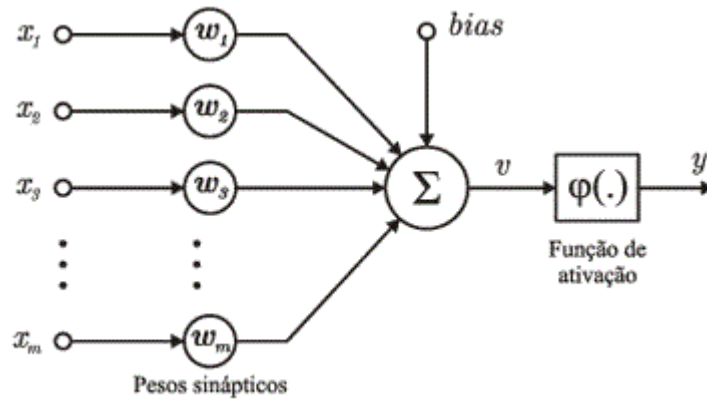


Figura 4. Modelo Matemático de um Neurônio Biológico

O primeiro índice da rede neural é sempre pré-definido e a sua multiplicação pelo respectivo peso representa o limiar de excitação [11].

$$net_i = \sum_{i=0}^n x_i \cdot w_i \quad (2.4)$$

$$y = f(net_i) \quad (2.5)$$

2.3 Multilayer Perceptron

O *Multilayer Perceptron* é uma generalização da rede Perceptron. É uma rede composta por uma camada de entrada, uma ou mais camadas intermediárias, também chamadas de camadas escondidas, e uma camada de saída (Figura 5). Em cada uma dessas camadas é possível conter um ou mais neurônios.

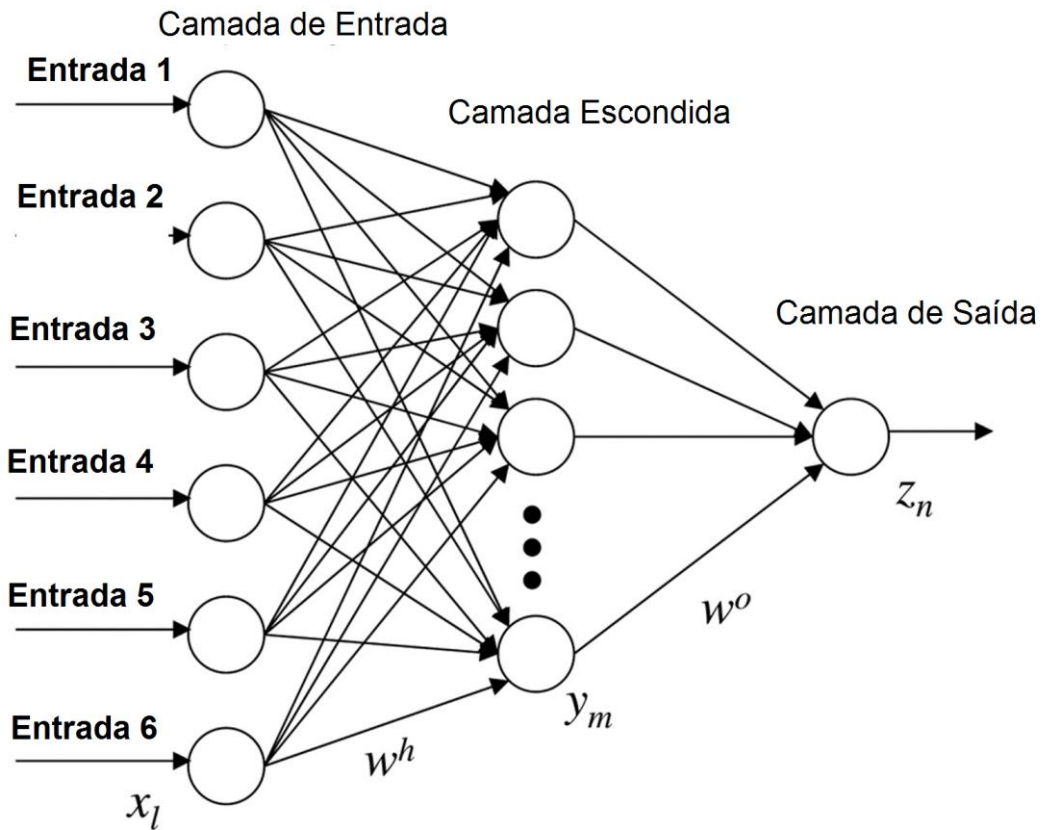
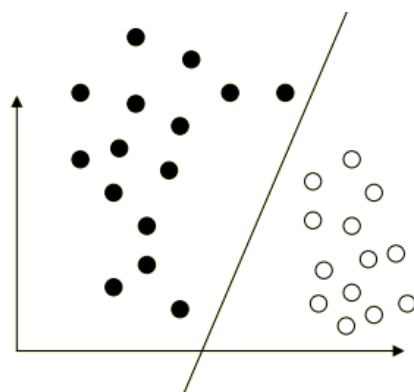
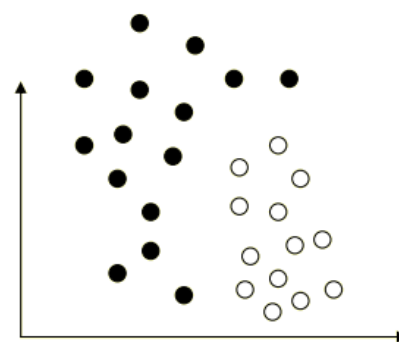


Figura 5. Arquitetura da Rede MLP

A adição de camadas permitiu que a rede MLP conseguisse resolver problemas não linearmente separáveis, pois permite a aproximação de qualquer função [11]. Um problema é linearmente separável se existir uma reta ou plano capaz de dividir os exemplos em diferentes classes, como na Figura 6, a seguir.



Linearmente separável



Não linearmente separável

Figura 6. Conjuntos linearmente separável e não linearmente separável

A rede MLP é do tipo *feedforward*, ou seja, a informação se propaga só em uma direção, sem ciclos, da camada de entrada para as camadas escondidas e, por fim, para a camada de saída.

Contudo, quanto mais camadas escondidas são adicionadas, mais complexo torna-se o treinamento. Ajustes de peso da camada de entrada e das camadas escondidas se tornam cada vez menos triviais. O mais difundido dos algoritmos de treinamento, no caso da MLP, é o *Backpropagation* [11].

2.3.1 Algoritmo Backpropagation

O treinamento é um conjunto de regras adicionadas às RNAs com o intuito de aumentar a sua capacidade de generalização. Ao ajustar os pesos das conexões entre os neurônios é possível obter um modelo com capacidade de generalização. Os treinamentos podem ser classificados como: aprendizado supervisionado, quando existe uma resposta para o problema dado de forma que seja possível comparar com a resposta obtida, aprendizado não supervisionado, quando não existe agente externo indicando a resposta desejada para os padrões de entrada e aprendizado por reforço, quando um agente externo avalia a resposta fornecida pela rede.

Backpropagation é o algoritmo de treinamento supervisionado que busca minimizar o erro do aprendizado, se utilizando do método do gradiente descendente, realizando correção dos pesos no sentido contrário à propagação da informação. Ou seja, da camada de saída para a camada escondida, e das camadas escondidas até a camada de entrada.

Este algoritmo possui duas fases: fase de propagação do sinal (conhecida por *Fase Forward*), e Fase do ajuste dos pesos (também chamada de *Fase Backward*).

Há, ainda, dois parâmetros importantes para o algoritmo. A taxa de aprendizado (α) e o momentum (β).

A taxa de aprendizado corresponde ao tamanho do passo em direção à correção do erro. Se o valor da taxa de aprendizado for muito baixo o treinamento se torna lento bem como o mesmo pode convergir prematuramente para mínimos

indesejáveis, porém se o valor for muito alto o processo de aprendizagem pode não ser realizado de forma adequada e a rede não terá boa capacidade de generalização.

O objetivo do momentum é aumentar a velocidade do treinamento da rede através da aceleração da convergência e diminuir o risco de ficar preso a mínimos locais [12].

Durante a fase *forward* ocorre a propagação do sinal da camada entrada até a camada de saída, de acordo com as equações (2.4) e (2.5). Ao final, a saída é calculada.

Já na fase *backward* ocorre a propagação do erro no sentido contrário, da camada de saída até a camada de entrada.

Como o aprendizado é supervisionado é possível comparar a saída obtida na fase *forward* à saída desejada do conjunto de dados correspondente às entradas utilizadas. Para obterem-se os novos valores dos pesos é necessário calcular as sensibilidades de cada neurônio.

A sensibilidade δ_j para a camada de saída é dada pela equação:

$$\delta_j = (d_j - y_j) f'(net_j) \quad (2.6)$$

Onde d_j corresponde à saída desejada para as entradas fornecidas, e y_j representa a saída obtida pelo mesmo neurônio com as mesmas entradas, e $f'(net_j)$ é a derivada da função de ativação do neurônio de saída.

Para as outras camadas a sensibilidade δ_j para cada neurônio j da respectiva camada, é calculado de acordo com a equação:

$$\delta_j = f'(net_j) \sum_i \delta_i w_{i,j} \quad (2.7)$$

Onde $f'(net_j)$ representa a derivada da função de ativação do neurônio corrente, δ_i corresponde à sensibilidade propagada pelo i -ésimo neurônio da camada logo à frente, e $w_{i,j}$ é o peso j do neurônio i .

O ajuste de pesos a cada iteração t é dado pela equação:

$$\Delta w_{i,j}(t+1) = w_{i,j}(t) + \alpha \delta_j x_i + \beta (w_{i,j}(t) - w_{i,j}(t-1)) \quad (2.8)$$

Onde $\Delta w_{i,j}(t+1)$ corresponde ao novo valor atribuído ao i -ésimo peso do neurônio j , $w_{i,j}(t)$ representa o valor do i -ésimo peso do neurônio j no instante (iteração) t , α é a taxa de aprendizado, δ_j é a sensibilidade, x_i é o valor do sinal de neurônio, β é o momentum e $w_{i,j}(t-1)$ é o valor do peso de índice i do neurônio j no instante $t-1$.

2.3.2 Critério de Parada do Treinamento da Rede

Determinar o número exato de iterações é um grande desafio no treinamento das redes MLP uma vez que um número alto de iterações pode causar superajustamento (*overfitting*) fazendo com que a rede decore os resultados e perca a capacidade de generalização. Por outro lado o baixo número de iterações pode causar problemas na convergência da rede [12].

Uma das alternativas para resolver o problema da parada do treinamento é a técnica da validação cruzada. Ao invés de definir o número exato de iterações de ajuste de pesos no treinamento, divide-se conjunto de dados em 3 partes: treinamento, validação e testes. A porcentagem da divisão varia, mas geralmente é

de 50%, 25% e 25%, respectivamente [12]. Para cada iteração do *backpropagation*, com os pesos ajustados, a rede é testada com o subconjunto de validação e o erro da previsão é calculado ao fim de cada iteração.

Enquanto o erro da validação decresce, a rede está generalizando. A partir do momento em que o erro começa a crescer juntamente com o erro do treinamento decrescendo, a rede começa a decorar as entradas, perdendo a capacidade de generalização. Nesse momento deve-se parar o treinamento.

O gráfico do melhor momento para parar o treinamento da rede é mostrado a seguir:

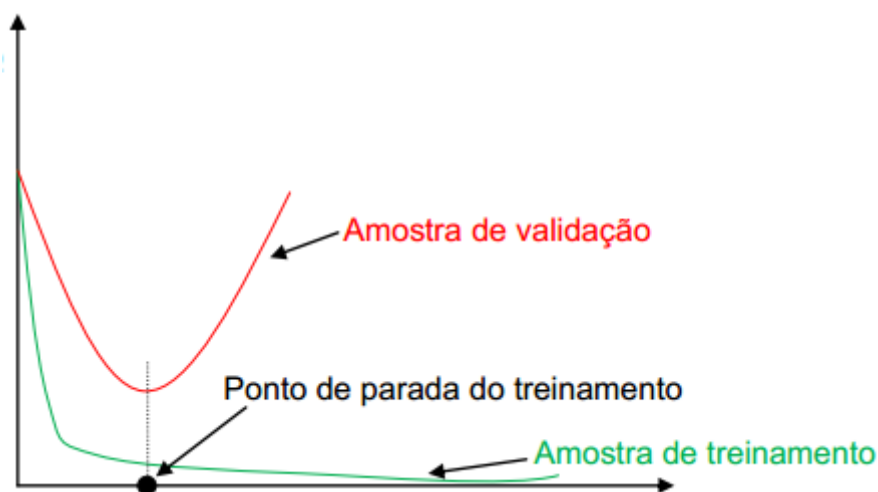


Figura 7. Gráfico do melhor ponto de parada do treinamento

2.4 Redes Neurais Recorrentes de Elman

Redes neurais recorrentes, do inglês *recurrent neural network* (RNN), são sistemas dinâmicos com propriedades sensitivas a sequencias temporais, onde as ligações entre as unidades formam um ciclo de instruções. Isto cria um estado interno da rede que permite que exiba comportamento temporal dinâmico. Ao contrário das redes neurais progressivas, RNNs podem usar sua memória interna para processar as sequências arbitrárias de insumos. Isso as tornam aplicáveis a tarefas como reconhecimento de escrita, onde eles têm conseguido os melhores resultados conhecidos [19].

2.4.1 Topologia da Rede

Nas Redes de Elman, além das unidades de entrada, intermediária e de saída há também unidades de contexto, como nas redes recorrentes em geral. As unidades de entrada e de saída interagem com o ambiente externo, enquanto as unidades intermediárias e de contexto não o fazem.

As unidades de entrada são apenas unidades de armazenamento (*buffer*) que passam os sinais sem modificá-los.

As unidades de saída são unidades lineares que somam os sinais que recebem. Neste trabalho, as unidades intermediárias têm funções de ativação não lineares, embora, segundo a teoria, possam ser lineares.

As unidades de contexto são usadas apenas para memorizar as ativações anteriores das unidades intermediárias e podem ser consideradas como atraso no tempo em um passo.

As conexões *feedforward* são modificáveis, e as conexões recorrentes são fixas, motivo pelo qual a RE é apenas parcialmente recorrente.

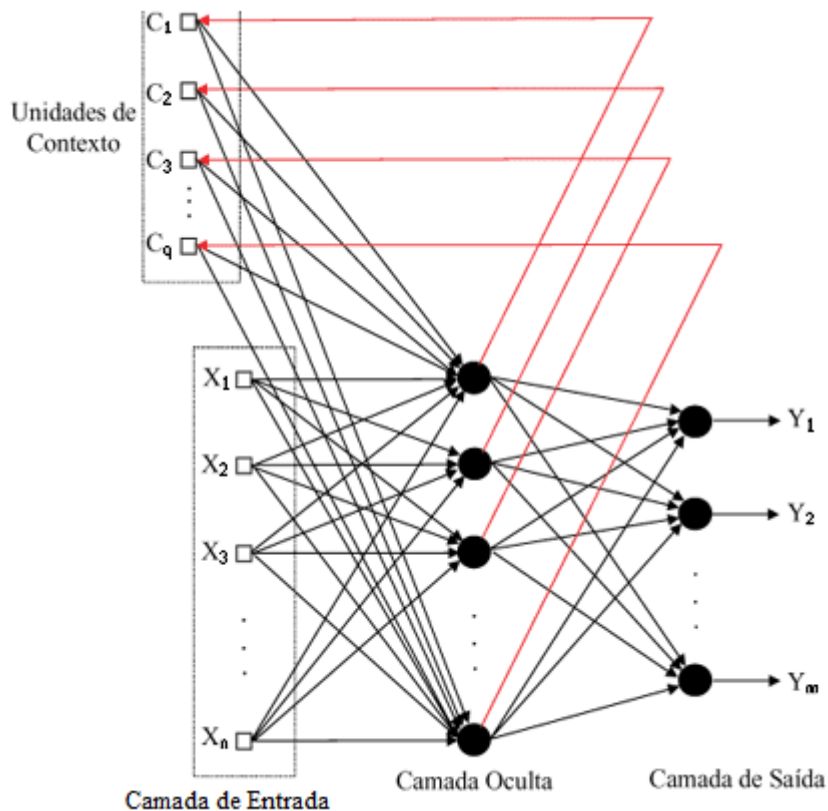


Figura 8. Topologia da Rede de Elman

2.4.2 Treinamento da Rede de Elman

Em um intervalo de tempo específico k , as ativações das unidades intermediárias (em $k-1$) e as entradas correntes (em k) são utilizadas como entradas da rede. Em um primeiro estágio, *feedforward*, estas entradas são propagadas para frente para produzir as saídas. Posteriormente, a rede é treinada com o algoritmo de aprendizagem *backpropagation* padrão. Após este passo de treinamento, as ativações das unidades intermediárias no tempo k são introduzidas, através das ligações recorrentes nas unidades de contexto, sendo salvas nestas unidades para o próximo passo de treinamento ($k+1$).

2.5 Redes Neurais Recorrentes de Jordan

Assim como nas redes de Elman, as redes neurais recorrentes de Jordan também possuem uma camada de contexto. Porém, nas redes de Jordan esta camada de contexto é realimentada pela camada de saída, ao invés da camada escondida. Esta camada de contexto continuará alimentando a camada escondida, juntamente com a camada de entrada. Em relação a quantidade de neurônios na camada de contexto, também será a mesma quantidade da camada escondida, assim como na rede de Elman.

Uma outra diferença entre as redes de Jordan e de Elman é que a camada de contexto das redes de Jordan também é realimentada por ela própria, como pode ser visto na Figura 9, a seguir.

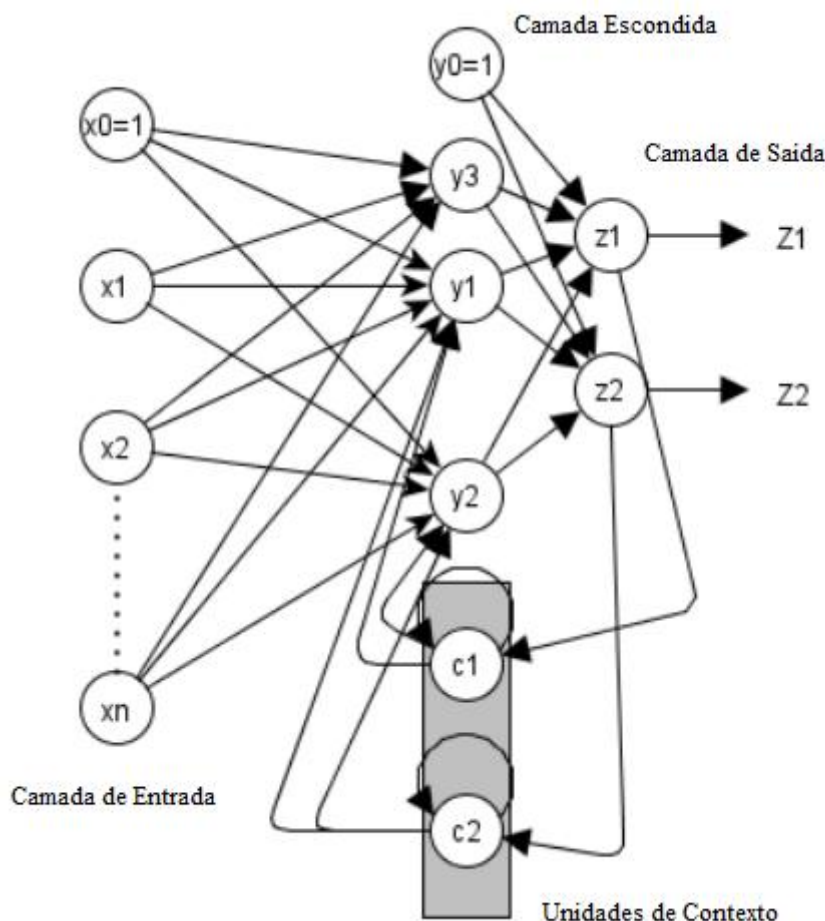


Figura 9. Topologia da Rede de Jordan

2.6 Encog

Encog [20] [21] é um *framework* que implementa vários tipos de RNA. Ele faz parte de um projeto *open source* e fornece as ferramentas necessárias para se utilizar todos estes tipos de redes neurais.

A ferramenta está disponibilizada nas linguagens de programação Java, C# e C++. Também possui uma interface gráfica, para uma implementação mais intuitiva das redes, sem precisar usar programação, que é implementada em Java,

Além da base para implementação das RNAs, o Encog também conta com suporte para tratamento de dados, funções de ativação e algoritmos de treinamento. Além disso, provê algumas outras técnicas de aprendizado, como algoritmos genéticos, modelo de Markov, entre outros.

O Encog também dá suporte ao critério de parada do treinamento, porém deixa a desejar neste ponto, pois o treinamento é parado apenas por número de iterações ou taxa de erro.

Capítulo 3

Metodologia

Este capítulo demonstra como as fundamentações teóricas serão utilizadas para que o objetivo do projeto seja alcançado. Os algoritmos de redes neurais MLP e Elman e Jordan que foram utilizados pertencem ao Encog, também descrito no capítulo de fundamentação teórica.

3.1 Base de Dados

A base de dados utilizada foi uma base de dados real cedida, a qual dar-se-á o nome de parque A. Os dados coletados na base abrangem o período de 1º de dezembro de 2011 à 31 de julho de 2012, com valores de trinta em trinta minutos. Os dados que estão contidos na base são: ano, mês, dia, hora, minutos e velocidade do vento medida.

3.2 Pré-Processamento dos Dados

A primeira etapa do pré-processamento foi à eliminação de entradas cujo valor da velocidade do vento era vazio. Na base de dados não foi encontrada nenhuma ocorrência de entradas de valor vazio.

Depois disso foi realizada uma busca por valores não existentes no universo de análise de velocidade dos ventos, como velocidade negativa ou maior do que 110 km/h. Por haver pouquíssimos valores desconsiderados em relação à quantidade de dados na base, a eliminação foi realizada sem problemas.

3.3 Normalização dos Dados

O próximo passo é a normalização dos dados. A normalização serve para evitar que grandes variações dos valores da entrada dificultem o treinamento e o aprendizado da rede, além de evitar que haja ineficiência do algoritmo de treinamento. Esta também servirá para manter os valores das entradas proporcionais aos limites das funções de ativação que serão utilizadas na rede. Se a função de ativação da rede neural for a sigmoide logística (o que se aplica a este trabalho), seus valores serão limitados pelo intervalo [0 e 1]. Logo os dados serão geralmente normalizados entre [0,10 e 0,90] ou [0,15 e 0,85], devido ao algoritmo *backpropagation* se tornar lento quando a derivada da função de ativação fica próxima de zero [12]. Neste trabalho será utilizado o intervalo de [0,15 e 0,85]. A equação a seguir demonstra como é feita a normalização de dados:

$$y = \frac{(b - a) \times (x_i - x_{\min})}{x_{\max} - x_{\min}} + a \quad (3.1)$$

Onde y é o valor normalizado, b é o intervalo máximo da normalização (neste caso 0,85), a é o intervalo mínimo da normalização (neste trabalho 0,15), x_i é o valor da entrada, x_{\min} é o valor mínimo das variáveis de entrada e x_{\max} é o valor máximo das variáveis de entrada.

3.4 Conjuntos de Dados

O conjunto de dados, devido ao uso da validação cruzada para o treinamento das RNAs, foi dividido em três partes.

A primeira metade dos dados foi utilizada para o treinamento da rede. Este primeiro conjunto é responsável pelo aprendizado e pelo ajuste dos pesos.

A outra metade foi dividida ao meio para o conjunto de validação e testes. O segundo conjunto, utilizado para validação, ficará responsável por indicar o momento ideal para a parada do treinamento da rede. Por sua vez, o terceiro conjunto, utilizado para testes, ficará responsável por avaliar o desempenho da rede com os dados restantes.

Como neste trabalho se usa apenas uma variável para as entradas (velocidade do vento) os dados foram dispostos sempre em ordem. Além disso, as redes neurais recorrentes (neste trabalho, as redes de Elman e Jordan) requerem que os dados estejam dispostos sempre em ordem, visto que os utilizam para realimentar a camada escondida.

3.5 Desnormalização dos Dados

Por trabalhar com os dados normalizados, os resultados da RNA também serão normalizados. Para descobrirmos o valor real que foi previsto, devemos executar a função inversa da normalização, que é dada por:

$$x_i = \frac{(y - a) \times (x_{\max} - x_{\min})}{b - a} + x_{\min} \quad (3.2)$$

Onde x_i é o valor da entrada (desnormalizado), y é o valor normalizado, b é o intervalo máximo da normalização (neste caso 0,85), a é o intervalo mínimo da normalização (neste trabalho 0,15), x_{\min} é o valor mínimo das variáveis de entrada e x_{\max} é o valor máximo das variáveis de entrada.

3.6 Implementação das RNAs

Embora seja uma rede neural muito utilizada em pesquisas, o MLP necessita que vários de seus parâmetros sejam configurados e a escolha de cada um destes parâmetros influencia diretamente no resultado final da previsão.

A seguir, seguem os principais parâmetros do MLP e do algoritmo de treinamento *backpropagation*:

- Número de neurônios na camada de entrada;
- Número de neurônios na camada escondida (apenas uma camada foi utilizada neste trabalho);
- Número de neurônios na camada de saída;
- Número de neurônios na camada de contexto (apenas para as redes de Elman e Jordan);
- Função de ativação;
- Taxa de Aprendizado;
- Momentum;
- Critério de Parada.

3.6.1 Camada de Entrada

Para definir as camadas de entrada, foi feito um estudo de sensibilidade com simulações utilizando três, seis e doze horas defasadas para a realização da previsão, utilizando seis, doze e vinte e quatro (24) neurônios, respectivamente.

3.6.2 Camada Escondida

Feitas diversas simulações, verificou-se que o melhor ajuste para a camada escondida ocorreu para dezoito neurônios.

3.6.3 Camada de Saída

Como se quer prever às seis horas seguintes, e a base mostra medições de trinta em trinta minutos, serão utilizados doze neurônios na camada de saída.

3.6.4 Camada de Contexto

As redes recorrentes de Elman e Jordan se utilizam também da Camada de Contexto, contêm a mesma quantidade de neurônios da camada escondida. Esta camada servirá para realimentação da camada escondida a cada ciclo de treinamento da rede.

3.6.5 Função de ativação

Na transposição dos dados de camada para camada é utilizada uma função de ativação. Esta corresponde a um limiar que restringe a propagação do impulso nervoso à transposição de certo nível de atividade, mapeando o potencial da camada anterior sobre a camada atual. A sigmoide logística, equação 3.3, é a mais utilizada. É uma função real cujo domínio é o conjunto dos números reais, a derivada é sempre positiva e a faixa de valores é limitada.

$$y = f(\textit{net}) = \frac{1}{1 + e^{-\textit{net}}} \quad (3.3)$$

3.6.6 Taxa de Aprendizado

A taxa de aprendizado é uma constante de proporcionalidade no intervalo [0,1]. Entretanto, a alteração do valor destas constantes gerará uma mudança nos pesos. Quanto maior o valor do gradiente, maior será a velocidade do aprendizado, o que pode levar à uma oscilação do modelo na superfície de erro. O ideal seria utilizar a maior taxa de aprendizado possível que não levasse à uma oscilação, resultando em um aprendizado mais rápido. O treinamento das redes com *backpropagation* pode demandar muitos passos no conjunto de treinamento, resultando um tempo de treinamento consideravelmente longo. Se for encontrado um mínimo local, o erro para o conjunto de treinamento para de diminuir e estaciona em um valor maior que o aceitável.

Por meio de realização de testes, a taxa de aprendizado utilizada neste trabalho foi de valor 0,3.

3.6.7 Momentum

Uma maneira de aumentar a taxa de aprendizado sem levar à oscilação é incluir o termo *momentum*, uma constante que determina o efeito das mudanças passadas dos pesos na direção atual do movimento no espaço de pesos. Desta forma, o *momentum* leva em consideração o efeito de mudanças anteriores de pesos na direção do movimento atual no espaço de pesos. O *momentum* torna-se útil em espaços de erro que contenham longas gargantas, com curvas acentuadas ou vales com descidas suaves.

Por meio de realização de testes, o *momentum* utilizado neste trabalho foi de valor 0,7.

3.6.8 Critério de Parada

O critério de parada utilizado no trabalho foi o mesmo descrito na seção 2.3.2. Quando o erro alcança um mínimo no conjunto de validação, e este conjunto de validação torna a crescer, o treinamento é parado naquele momento, a fim de evitar *underfitting* (por falta de treinamento) e *overfitting* (excesso de treinamento e perda da capacidade de generalização).

3.6.9 Treinamento da RNA

O algoritmo de treinamento utilizado nas RNAs foi o *backpropagation*, descrito na seção 2.3.1.

3.7 Teste Comparativo

Para realizar a comparação de desempenho das arquiteturas de RNA para a previsão da velocidade do vento, foi utilizado o erro médio quadrático (EMQ). Para

tal comparação foi utilizado um método estatístico. O nível de significância estabelecido nos testes foi de 5% (0,05).

Para a realização dos testes comparativos foi aplicado o teste não paramétrico de Wilcoxon [20] [21].

Os métodos estatísticos paramétricos assumem que os dados pertencem a alguma distribuição e que por causa disso podem ser descritos em poucos parâmetros, que caracterizam essa distribuição. [20] Já os testes não paramétricos, não fazem qualquer tipo de afirmação sobre a natureza da distribuição dos dados, podendo ser utilizados quando esta distribuição é desconhecida, ou não é possível o uso de técnicas paramétricas [21].

3.7.1 Teste de Igualdade

O teste de Wilcoxon verifica a magnitude da diferença entre dois dados através da comparação das medianas das amostras. Esse teste assume que a distribuição das diferenças é simétrica.

O primeiro passo da comparação foi verificar se as duas amostras comparadas são iguais ou diferentes. Para isso, foram elaboradas as seguintes hipóteses:

- H_0 : não existe diferença entre as medianas dos erros das amostras;
- H_1 : as medianas das duas amostras são diferentes.

3.7.2 Comparação entre as RNAs

A avaliação da previsão das arquiteturas das RNAs ocorreu através da comparação do EMQ da previsão. A arquitetura que obteve o menor EMQ foi considerada a melhor rede.

Para realizar a comparação, os desempenhos das arquiteturas são comparados aos pares. Cada rede MLP foi comparada com todas as outras redes,

sejam elas também MLP, redes de Elman ou redes de Jordan também seguindo o mesmo procedimento para cada rede de Elman e para cada rede de Jordan.

O teste de Wilcoxon foi aplicado para avaliar estas duas hipóteses para cada par de arquiteturas de RNA. Quando a H_1 é obtida como resultado, considera-se que a melhor arquitetura é a que apresentar o menor EMQ entre as duas do par avaliado.

3.7.3 Comparação com o Modelo de Persistência

Após realizada a comparação entre as RNAs, são eleitas as melhores redes de cada arquitetura, e estas serão avaliadas pelo teste de Wilcoxon.

Ao final, apenas a melhor RNA será comparada com o modelo de persistência, também com a avaliação do teste de Wilcoxon.

Capítulo 4

Resultados

Este capítulo apresenta os resultados dos passos realizados no capítulo 3.

4.1 Estudo de Sensibilidade da Camada de Entrada

Para avaliar o desempenho de uma RNA se utiliza um conjunto de dados que não fez parte do processo de treinamento, também conhecido como conjunto de testes. No final do treinamento, este conjunto é apresentado para a RNA e os valores dos erros de previsão são obtidos a partir da seguinte equação:

$$EMQ = \frac{1}{n} \sum_{i=1}^n (P_i - O_i)^2 \quad (4.1)$$

Onde n é a quantidade de valores do conjunto de testes, P_i é o i -ésimo valor previsto pela RNA, e O_i é o i -ésimo valor ocorrido.

Tendo em vista o objetivo de realizar um teste comparativo estatisticamente confiável, o processo de treinamento e avaliação é repetido trinta vezes, para todas as RNAs propostas.

A seguir, as tabelas 4.1, 4.2 e 4.3 mostram os EMQs das redes neurais MLP, Elman e Jordan.

Tabela 4.1 Erros Médios Quadráticos para a rede MLP

Amostra	MLP 6 Entradas	MLP 12 Entradas	MLP 24 Entradas
1	0,0128112438906673	0,0123942073242801	0,0097031472040919
2	0,0163311205943120	0,0231614679619551	0,0086430639820220
3	0,0104743776337600	0,0099024016288040	0,0090392022448825
4	0,0101405625937468	0,0190972159094561	0,0179854388146545
5	0,0150387326994937	0,0157829784423020	0,0184326660263287
6	0,0110161542878433	0,0108594059401485	0,0101057461335325
7	0,0112723469355080	0,0114244225920716	0,0080088914499614
8	0,0201110614449429	0,0128157242625291	0,0081665121390169
9	0,0110880262172097	0,0116704256778621	0,0106834602522582
10	0,0137957934422167	0,0118973134900318	0,0103377012346318
11	0,0114660311877945	0,0119358404094125	0,0148710211637221
12	0,0098844282936844	0,0161476491879153	0,0110095875464055
13	0,0092446319071804	0,0172900223303555	0,0116627550792352
14	0,0193716657837937	0,0118660361170952	0,0104087554316878
15	0,0183176138848241	0,0115246040614652	0,0122399101616826
16	0,0106927388325450	0,0109886004039122	0,0113611467878004
17	0,0085996485232519	0,0126764498376119	0,0159880085357039
18	0,0105021368004522	0,0168495455514720	0,0131212094570318
19	0,0119440838069100	0,0122415173107857	0,0077507627275667
20	0,0098281761225737	0,0121750472094682	0,0095230061929504
21	0,0175858314660568	0,0104779148568055	0,0090014648775597
22	0,0103242704549781	0,0118940401378862	0,0108334128525007
23	0,0136800425755897	0,0101314549541634	0,0133325466661724
24	0,0258906823980112	0,0096818560119806	0,0156875407073338
25	0,0119596584032306	0,0132105082139101	0,0104050562607518
26	0,0151166370646287	0,0163061770098272	0,0153745631759259
27	0,0106789928088256	0,0135278615544199	0,0106979578132868
28	0,0150842002656158	0,0098792051216682	0,0143330239744901
29	0,0179626292885116	0,0121069904226269	0,0108506076252649
30	0,0107325161519656	0,0109509919399825	0,0165308525240036

Tabela 4.2 Erros Médios Quadráticos para a rede de Elman

Amostra	Elman 6 Entradas	Elman 12 Entradas	Elman 24 Entradas
1	0,0129930858463378	0,0188804503035254	0,0211126344562698
2	0,0146792349512798	0,0133944121608248	0,0149696756285406
3	0,0145790943185329	0,0121799077404567	0,0180651922943465
4	0,0141111622163621	0,0152600805088657	0,0135898434149386
5	0,0131087819332304	0,0288008984711588	0,0207116315410947
6	0,0137776110024561	0,0140882371906129	0,0152237258924866
7	0,0133244123409875	0,0162679360151194	0,0098059094799456
8	0,0127499772877463	0,0146135991595447	0,0238961130375215
9	0,0137043486918525	0,0141322602187875	0,0173139250325040
10	0,0164085980907864	0,0176528648205865	0,0209332495193618
11	0,0156916854339769	0,0121594449318415	0,0143020120337013
12	0,0244416544724605	0,0198879590179976	0,0150551501825935
13	0,0116688510352785	0,0145943403049479	0,0154730153944668
14	0,0205388023752813	0,0166056103913398	0,0120070380218788
15	0,0145456667986918	0,0134481513721303	0,0166975815855626
16	0,0113032773223653	0,0109121079735590	0,0133528341468866
17	0,0161820756144367	0,0200921212572835	0,0097023046443698
18	0,0119179498444918	0,0188586045919167	0,0102547784085334
19	0,0132979239300843	0,0190768071346061	0,0102152814374936
20	0,0122621549277585	0,0104670774911430	0,0268770351619172
21	0,0148750558688917	0,0140592082813134	0,0125073076440752
22	0,0223764429602269	0,0110562371199983	0,0124327940176728
23	0,0149676894690506	0,0194157504099076	0,0228131700455113
24	0,0125608465551986	0,0178012551336848	0,0172381715836899
25	0,0127523583550922	0,0101137932972707	0,0175791932515027
26	0,0154382867613469	0,0160903282435058	0,0154034186988566
27	0,0122756567840928	0,0210859717035537	0,0224735618831087
28	0,0140510171458556	0,0213557077844395	0,0225867351183532
29	0,0175548217161382	0,0141374402368407	0,0135507669210376
30	0,0163445118299934	0,0179263503840732	0,0107491021187413

Tabela 4.3 Erros Médios Quadráticos para a rede de Jordan

Amostra	Jordan 6 Entradas	Jordan 12 Entradas	Jordan 24 Entradas
1	0,01367769467006170	0,01616667040063360	0,01470227920203360
2	0,01348962772853720	0,01304288245077660	0,01248022700553820
3	0,01589060976470470	0,01052103250333950	0,01201287301200020
4	0,01271929041530150	0,01027118193550880	0,01253404206603140
5	0,01341593673921940	0,01997527592216160	0,01256179245186440
6	0,01533627558845860	0,01352749273686530	0,01094336235554200
7	0,01571924793014390	0,01061186439916580	0,01030275936190950
8	0,02463237874935830	0,01026850024644430	0,00957698943181413
9	0,01999138309414510	0,01260401465057080	0,01418245495846560
10	0,01269511334759460	0,01088284062927300	0,02324861200019120
11	0,01292263253650550	0,01341518026848550	0,01261754349203830
12	0,00935706824001906	0,01778403076873390	0,01181578607238890
13	0,01880757741465420	0,01871980131014140	0,01031647117502690
14	0,00944876085965102	0,01256067427445990	0,01038608789152590
15	0,01207099312247950	0,01099075416014880	0,01496549508978850
16	0,01540721926523720	0,01323248100638410	0,01119965615284650
17	0,01085314346056250	0,02285137142900910	0,01324601936026540
18	0,01467226606628440	0,01004002569395900	0,00990671758220378
19	0,01215156918021890	0,01267569134993550	0,02512483586592290
20	0,01116585706766380	0,00975536811950590	0,01004976753493120
21	0,01228374807233150	0,01338345784347150	0,01060379060957070
22	0,01218089009556930	0,01177214446460260	0,01373833616938970
23	0,01063483856332200	0,01371720916405960	0,01656067207416770
24	0,01779179676660640	0,01291091290079480	0,01534773425996270
25	0,01207748543537180	0,01288035725380620	0,01606526242780340
26	0,01252346437675090	0,01146805450534020	0,01295215714949850
27	0,01208858444795630	0,01691599120337730	0,01819331638183160
28	0,01036816112357220	0,02196241482282850	0,01945328146364570
29	0,01718394686998460	0,02069478199332040	0,01028828914536210
30	0,01347184770642210	0,01796614339094830	0,01408249803333620

Pode-se observar na tabela 4.4 a média dos EMQs apresentados nas tabelas 4.1, 4.2 e 4.3. Essas médias serão utilizadas para definir quais as redes serão melhores nos estudos comparativos, se forem estatisticamente desiguais.

Tabela 4.4 Média dos erros médios quadráticos de todas as RNAs

	Média dos EMQs
MLP 6 Entradas	0,0133648678586708
MLP 12 Entradas	0,0130289291957402
MLP 24 Entradas	0,0118696339680819
Elman 6 Entradas	0,0148161011960095
Elman 12 Entradas	0,0161471637883612
Elman 24 Entradas	0,0162297717532321
Jordan 6 Entradas	0,0138343136232896
Jordan 12 Entradas	0,0141189533932684
Jordan 24 Entradas	0,0136486369925632

O estudo de sensibilidade foi realizado nas RNAs para avaliar a quantidade de entradas que seria melhor para cada rede neural. O teste de Wilcoxon foi feito par a par entre todas as arquiteturas de cada RNA. A seguir, as tabelas 4.5, 4.6 e 4.7 mostram a avaliação do teste de Wilcoxon para cada RNA.

Tabela 4.5 Comparação das entradas nas redes MLP

	MLP 6 Entradas	MLP 12 Entradas	MLP 24 Entradas
MLP 6 Entradas	X	0,786	0,1381
MLP 12 Entradas	X	X	0,05512
MLP 24 Entradas	X	X	X

Tabela 4.6 Comparação das entradas nas redes de Elman

	Elman 6 Entradas	Elman 12 Entradas	Elman 24 Entradas
Elman 6 Entradas	X	0,1381	0,2244
Elman 12 Entradas	X	X	0,959
Elman 24 Entradas	X	X	X

Tabela 4.7 Comparação das entradas nas redes de Jordan

	Jordan 6 Entradas	Jordan 12 Entradas	Jordan 24 Entradas
Jordan 6 Entradas	X	0,9707	0,5819
Jordan 12 Entradas	X	X	0,6022
Jordan 24 Entradas	X	X	X

Analisando as tabelas, vê-se que todas as entradas, de acordo com a avaliação do teste de Wilcoxon, são estatisticamente semelhantes, o que leva a poder usar qualquer quantidade de entradas (entre seis, doze e vinte e quatro) na comparação com o modelo de persistência.

4.2 Escolha da Quantidade de Neurônios na Camada Escondida

Para se implementar a camada escondida, foram realizadas simulações com seis, doze, dezoito, vinte e quatro, e trinta e seis neurônios. A figura abaixo mostra como a média dos EMQs se comportou em cada uma das simulações.

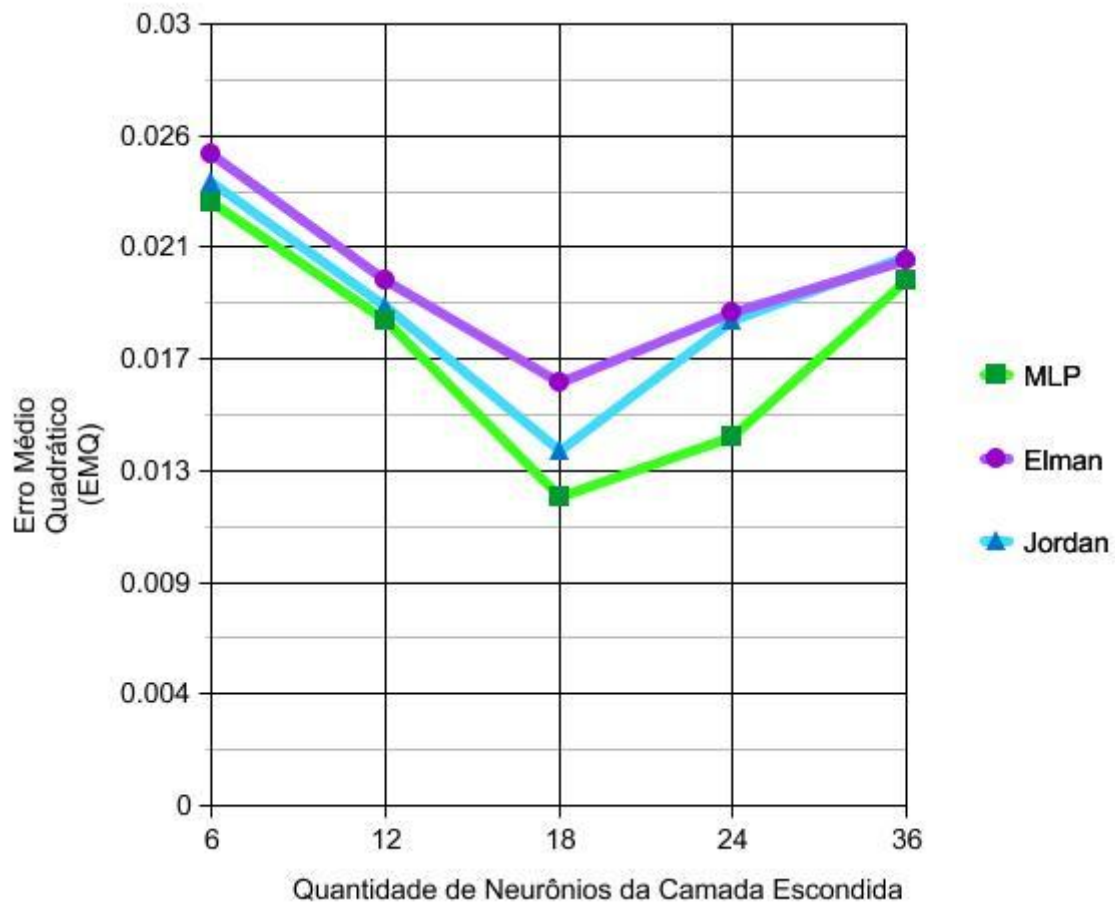


Figura 10. Simulações da camada escondida

Conclui-se que a camada de 18 (dezoito) neurônios apresenta melhores resultados, sendo escolhida para ser utilizada nas implementações das RNAs.

4.3 Testes Comparativos com o Modelo de Persistência

Para fazer a comparação entre as arquiteturas, serão utilizadas as redes que apresentaram melhor EMQ de cada uma delas. Escolheu-se a MLP de vinte e quatro entradas, a rede de Elman com seis entradas e por fim a rede de Jordan com vinte e

quatro entradas. A tabela 4.8 apresenta os resultados das avaliações dos testes de comparação.

Tabela 4.8 Comparação entre as arquiteturas das RNAs

	MLP 24 Entradas	Elman 6 Entradas	Jordan 24 Entradas
MLP 24 Entradas	X	0,00024200	0,06540000
Elman 6 Entradas	X	X	0,03997000
Jordan 24 Entradas	X	X	X

Analisando o resultado destes testes, pode-se concluir que a MLP e a rede de Jordan, apresentam melhores resultados que a rede de Elman. Já em relação a MLP e Jordan, quando comparadas, são estatisticamente semelhantes.

Já que se obtiveram duas redes estatisticamente idênticas perante o teste de Wilcoxon, estas duas redes serão comparadas com o modelo de persistência. Na tabela 4.9, são mostrados os resultados da comparação das redes com o modelo de persistência.

Tabela 4.9 Comparação das RNAs com o modelo de persistência

	MLP 24 Entradas	Jordan 24 Entradas
Modelo de Persistência	0,3470	0,01814

Analisando o resultado da tabela 4.9, verifica-se que o modelo de persistência apresentou um resultado estatístico significativo perante a rede de Jordan, por obter valor abaixo do nível de significância (0,05), e estatisticamente semelhante à rede MLP, por obter valor acima do nível de significância.

Capítulo 5

Considerações Finais

5.1 Conclusões

O crescimento da demanda por fontes de energia renováveis incentiva o desenvolvimento da energia eólica. A diminuição de seu custo de operação e instalação vem tornando essa alternativa energética cada vez mais viável. Porém, para que a energia eólica seja bem sucedida é necessária a previsão precisa da velocidade do vento.

Este trabalho de conclusão teve por principal objetivo aplicar técnicas de redes neurais artificiais conhecidas, na expectativa de comparar estas técnicas e suas arquiteturas propostas e, ao final, compará-las com o modelo de persistência, fazendo estimativas de velocidade do vento.

Para se alcançar este objetivo, foram implementadas redes neurais com diferentes técnicas e arquiteturas já conhecidas e, utilizando uma base de dados real, realizou-se diversas simulações e testes para posterior comparação dos resultados.

Através dos testes estatísticos foi comprovado que o desempenho dos modelos das redes MLP e recorrente de Jordan apresentaram um melhor desempenho do que a rede recorrente de Elman.

O modelo de persistência, por sua vez, mostrou desempenho estatisticamente superior à rede recorrente de Jordan, e um desempenho estatisticamente semelhante à rede MLP com 24 entradas.

5.2 Trabalhos Futuros

Para trabalhos futuros outras configurações de redes podem ser testadas, assim como:

- Utilizar outras variáveis de entrada, além da velocidade do vento.
- Definir novas configurações que causem alterações no EMQ
- Utilizar diferentes funções de ativação
- Utilizar outras bases de dados e realizar testes estatísticos.

Da mesma forma, outras técnicas de redes neurais podem ser utilizadas: como redes neurais adaptativas ou redes neurais construtivas.

Bibliografia

- [1] ALDABÓ, Ricardo. **Energia Eólica**. 1. ed. São Paulo: Editora ArtLiber. 2002. 156 p.
- [2] CAMARGO A. S. G., “**Análise da Operação das Usinas Eólicas de Camelinho e Palmas e Avaliação do Potencial Eólico de Localidades no Paraná**”.
- [3] CHESF-BRASCEP, 1987. Fontes Energéticas Brasileiras, Inventário/Tecnologia. Energia Eólica. V.1 De cata-ventos a aerogeradores: o uso do vento, Rio de Janeiro.
- [4] SÁ, A. L. **Atlas do Potencial Eólico Brasileiro**. DEWI Magazine, no. 19, p. 63-73, 2001.
- [5] TERCIOTE, Ricardo. **Eficiência Energética de um Sistema Eólico Isolado**. 7 f. Trabalho de Conclusão de Curso de Engenharia Mecânica, Universidade Estadual de Campinas, Campinas.
- [6] RODRIGUES, Guilherme. **Utilização de Redes Neurais para Previsão de Ventos no Horizonte de 24 Horas**. 2007. 119 f. Dissertação de Mestrado do Curso de Engenharia Elétrica, Universidade Federal do Rio de Janeiro, Rio de Janeiro.
- [7] SARAIVA, Rodrigo. **Localização óptima dos conversores de um parque eólico para minimização de custos totais**. 2007. 59 f. Dissertação de Mestrado do Curso de Engenharia Elétrica e de Computadores, Universidade Técnica de Lisboa, Lisboa.
- [8] BRAGA, A. de P.; PONCE DE LEON, F. d. C. A.; BERNARDA LUDERMIR, T. **Redes Neurais Artificiais: Teorias e Aplicações**. 2. ed. [S.l.]: LTC, 2000.
- [9] ELETROBRÁS. Proinfra. 2012. Acessado em 29 de setembro de 2013. Disponível em:
<<http://www.eletrobras.com/elb/data/Pages/LUMISABB61D26PTBRIE.htm>>.
- [10] FRANCISCO, W. de Cerqueira e. **Fontes de Energia**. 2008. Acessado em 01 de novembro de 2013. Disponível em:
<http://www.mundoeducacao.com.br/geogra_a/fontesenergia.htm>.

- [11] HAYKIN, S. **Redes Neurais: Princípios e Práticas**. 2. ed. [S.l.]: Bookman, 2007.
- [12] VALENÇA, Mêuser. **Fundamentos das redes Neurais**. 2. ed. rev. ampl. Pernambuco: Livro Rápido, 2009. 384 p.
- [13] VALENÇA, M. J. S. **Aplicando Redes Neurais: Um Guia Completo**. Recife: Livro Rápido, 2005.
- [14] TRIOLA, M. F. **Introdução a estatística: livros técnicos e científicos**. 7. ed. Rio de Janeiro: [s.n.], 1999. 410 p.
- [15] Makarov, Y, Hawkins, D., Leuze, E. and Vidov, J (2003). **California ISO Wind Generation Forecasting Service Design and Experience. Proceedings of American Wind Energy Association Conference**, Oregon, USA.
- [16] Siebert, N (2008). **Development of Methods for Regional Wind Power Forecasting**. p. 37-38, 2008.
- [17] Lei, M., Shiyang, L., Chuanwen, J., Hongling, L. and Yan, Z (2009). **A review on the forecasting of wind speed and generated power**. *Renew Sustain Energy Rev*, 13(4): 915–20.
- [18] Stefsos, A. **A review comparison of various forecasting techniques applied to mean hourly wind speed time series**. *Renew Energy*, 13(4): 915–20.
- [19] A. Graves, M. Liwicki, S. Fernandez, R. Bertolami, H. Bunke, J. Schmidhuber. **A Novel Connectionist System for Improved Unconstrained Handwriting Recognition**. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 5, 2009.
- [20] HEATON, J. **Introduction to Neural Networks With Encog 3 for C#**. 2 Ed. St.Louis, MO USA: Heaton Research, Inc, 2011.
- [21] HEATON, J. **Programming Neural Networks With Encog 3 In C#**. St. Louis, MO USA: Heaton Research, Inc, 2011.

[22] BOX, G. E. P.; HUNTER, J. S.; HUNTER, W. G. **Statistics for experimenters**. 2. ed. [S.I.]: Wiley-Interscience, 2005.

[23] GIBBONS, J. D.; CHAKRABORTI, S. **Nonparametric statistical Inference**. 4. Ed. New York: Marcel Dekker, 2003.

[24] ANEEL. **Atlas da Energia Eólica do Brasil**. 2012. Acessado em 21 de novembro de 2013. Disponível em: <<http://www.aneel.gov.br/>>.