

UTILIZAÇÃO DE REDES NEURAIS PARA ANÁLISE DE SÉRIES TEMPORAIS DE POTÊNCIA EÓLICA

Trabalho de Conclusão de Curso
Engenharia da Computação

Danilo da Silva Barbosa de Lima

Orientador: Prof. Dr. Mêuser Jorge Silva Valença

Danilo da Silva Barbosa de Lima

UTILIZAÇÃO DE REDES NEURAIS PARA ANÁLISE DE SÉRIES TEMPORAIS DE POTÊNCIA EÓLICA

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

Universidade de Pernambuco

Escola Politécnica de Pernambuco

Graduação em Engenharia de Computação

Orientador: Prof. Dr. Mêuser Jorge Silva Valença

Recife - PE, Brasil

24 de novembro de 2015

Declaro que revisei o Trabalho de Conclusão de Curso sob o título “*UTILIZAÇÃO DE REDES NEURAIS PARA ANÁLISE DE SÉRIES TEMPORAIS DE POTÊNCIA EÓLICA*”, de autoria de *Danilo da Silva Barbosa de Lima*, e que estou de acordo com a entrega do mesmo.

Recife, _____ / _____ / _____

Prof. Dr. Mêuser Jorge Silva Valença
Orientador

"Sempre que você andar 5 passos em direção ao horizonte, ele deve se afastar 5 passos. Se você andar 10 passos na direção dele, ele deve se afastar 10 passos, pois o horizonte não está lá para você alcançá-lo, mas para que você nunca pare de caminhar!"

Agradecimentos

Primeiro agradeço a Deus por me dar forças nos momentos em que precisei, para chegar até aqui e concluir este trabalho.

Agradeço ao professor Dr. Mêuser Valença que aceitou me orientar no desenvolvimento deste projeto. E todos que foram meus professores durante o curso.

Agradeço às alunas Sabrina Andrade e Larissa Soares que me auxiliaram no desenvolvimento do trabalho.

Agradeço à minha família por estar presente e me ajudar sempre que precisei.

Agradeço imensamente à minha mãe Sueli Maria da Silva por todo amor, carinho e dedicação.

Resumo

A expansão do consumo de energia tem aspectos negativos. Com o intuito de amenizar esses efeitos, a humanidade vem tentando diversificar a matriz energética com novas fontes que, além de não causarem danos ao ambiente, sejam renováveis. Entre essas novas fontes destaca-se a energia eólica. Para sua utilização é necessário lidar com algumas incertezas, tais como velocidade e direção do vento. Assim, este trabalho tem o objetivo de propor um sistema de previsão de potência eólica, utilizando em sua construção uma Rede Neural Artificial. Através de experimentos realizados com três topologias de redes neurais (MLP, RBF e SVM) e diferentes horizontes de entrada, verificou-se que a rede SVM com 48 variáveis se mostrou mais adequada para a previsão de potência eólica.

Palavras-chave: MLP, RBF, SVM, Energia Eólica, Redes Neurais Artificiais.

Abstract

The expansion of energy use has negative aspects. In order to mitigate these effects, humanity has been trying new sources of energy that do not cause damage to the natural environment and are renewable. Among these new sources stands out the wind. For this energy source to be used it is necessary to deal with some uncertainties, such as wind speed and direction. Thus, this work aims to propose a wind power forecasting system using an Artificial Neural Network. Through experiments conducted with three different neural network topologies (MLP, RBF and SVM) and different input horizons, it was verified that the SVM network with 48 variables was more appropriate suitable for wind power forecast.

Keywords: MLP, RBF, SVM, wind energy, Artificial neural networks.

Lista de ilustrações

| | |
|--|----|
| Figura 1 – Moinho de vento Persa | 5 |
| Figura 2 – Sail Windmill | 5 |
| Figura 3 – Post mill | 6 |
| Figura 4 – Turbina eólica americana | 6 |
| Figura 5 – Turbina eólica de Charles Brush | 7 |
| Figura 6 – Movimento das massas de ar | 8 |
| Figura 7 – Turbina de Darrieus de duas lâminas | 11 |
| Figura 8 – Turbina eólica de eixo horizontal | 11 |
| Figura 9 – Aerogerador com múltiplas pás | 12 |
| Figura 10 – Complementaridade entre geração hidrelétrica e eólica | 14 |
| Figura 11 – Potencial eólico brasileiro | 14 |
| Figura 12 – Neurônio Biológico | 16 |
| Figura 13 – Neurônio Artificial | 17 |
| Figura 14 – Arquitetura da rede MLP | 18 |
| Figura 15 – Fases do algoritmo <i>Backpropagation</i> | 18 |
| Figura 16 – Estrutura da rede RBF | 20 |
| Figura 17 – Hiperplano ótimo separando os dados com a máxima margem ρ | 22 |
| Figura 18 – Exemplos de valores e situações da variável de folga ξ | 23 |
| Figura 19 – Transformação de um problema não linearmente separável em um problema linearmente separável | 23 |
| Figura 20 – Comparação da Raiz do EMQ com 6 valores | 34 |
| Figura 21 – Comparação da Raiz do EMQ com 12 valores | 34 |
| Figura 22 – Comparação da Raiz do EMQ com 24 valores | 35 |
| Figura 23 – Comparação da Raiz do EMQ com 48 valores | 35 |

Lista de tabelas

| | |
|--|----|
| Tabela 1 – Representação de defasagem dos valores presentes na base de dados do Parque I | 26 |
| Tabela 2 – Representação da média das taxas de erro após 30 simulações com cada arquitetura utilizando 6 variáveis. | 36 |
| Tabela 3 – Resultados do teste da Soma dos Postos de <i>Wilcoxon</i> para 6 valores de entrada. | 36 |
| Tabela 4 – Representação da média das taxas de erro após 30 simulações com cada arquitetura utilizando 12 variáveis. | 37 |
| Tabela 5 – Resultados do teste da Soma dos Postos de <i>Wilcoxon</i> para 12 valores de entrada. | 37 |
| Tabela 6 – Representação da média das taxas de erro após 30 simulações com cada arquitetura utilizando 24 variáveis. | 38 |
| Tabela 7 – Resultados do teste da Soma dos Postos de <i>Wilcoxon</i> para 24 valores de entrada. | 38 |
| Tabela 8 – Representação da média das taxas de erro após 30 simulações com cada arquitetura utilizando 48 variáveis. | 38 |
| Tabela 9 – Resultados do teste da Soma dos Postos de <i>Wilcoxon</i> para 48 valores de entrada. | 38 |
| Tabela 10 – Representação das taxas de erro de cada rede escolhida como mais adequada para horizonte de entrada. | 39 |

Lista de abreviaturas e siglas

| | |
|---------|--|
| ANEEL | Agência Nacional de Energia Eólica |
| CBEE | Centro Brasileiro de Energia Eólica |
| EMA | <i>Mean absolute error</i> - Error Médio Absoluto |
| EMQ | <i>Mean squared error</i> - Erro Médio Quadrático |
| ERA | <i>Relative absolute error</i> - Erro Relativo Absoluto |
| EREC | <i>European Renewable Energy Council</i> - Conselho Europeu de Energia Renovável |
| GW | Gigawatts |
| KW | Quilowatts |
| MLP | <i>Multi-Layer Perceptron</i> |
| MW | Megawatts |
| ONS | Operador Nacional do Sistema Elétrico |
| PROINFA | Programa de Incentivo às Fontes Alternativas de Energia Elétrica |
| RBF | <i>Radial Basis Function</i> |
| REMQ | <i>Root mean squared error</i> - Raiz do Erro Médio Quadrático |
| RERQ | <i>Root relative squared error</i> - Raiz do Erro Relativo Quadrático |
| RNAs | Redes Neurais Artificiais |
| SIN | Sistema Elétrico Interligado Nacional |
| ST | Série Temporal |
| SVM | <i>Support Vector Machine</i> |
| VC | Vapnik e Chervonenkis |

Sumário

| | | |
|------------|---|-----------|
| 1 | INTRODUÇÃO | 1 |
| 1.1 | Motivação | 1 |
| 1.2 | Objetivos | 2 |
| 1.2.1 | Objetivos Gerais | 2 |
| 1.2.2 | Objetivos Específicos | 2 |
| 1.3 | Estrutura da Monografia | 3 |
| 2 | FUNDAMENTAÇÃO TEÓRICA | 4 |
| 2.1 | Energia Eólica | 4 |
| 2.1.1 | Histórico | 4 |
| 2.1.2 | Geração de Energia Eólica | 7 |
| 2.1.3 | Turbina Eólica | 9 |
| 2.1.3.1 | Turbinas Eólicas de Eixo Vertical | 10 |
| 2.1.3.2 | Turbinas Eólicas de Eixo Horizontal | 10 |
| 2.1.4 | Energia Eólica no Mundo | 10 |
| 2.1.5 | Energia Eólica no Brasil | 13 |
| 2.2 | Redes Neurais | 15 |
| 2.2.1 | Introdução | 15 |
| 2.2.1.1 | Histórico | 15 |
| 2.2.1.2 | Neurônio Biológico | 15 |
| 2.2.1.3 | Neurônio Artificial | 16 |
| 2.2.2 | Multi-Layer Perceptron | 16 |
| 2.2.3 | Radial Basis Function | 20 |
| 2.2.4 | Support Vector Machine | 22 |
| 3 | MÉTODOLOGIA | 25 |
| 3.1 | Base de Dados | 25 |
| 3.2 | Processamento dos Dados | 25 |
| 3.2.1 | Normalização dos Dados | 25 |
| 3.2.2 | Defasagem dos Dados | 26 |
| 3.3 | Arquiteturas de Redes Neurais na Previsão de Potência Eólica | 27 |
| 3.3.1 | <i>Multi-Layer Perceptron</i> | 27 |
| 3.3.2 | <i>Radial Basis Function</i> | 28 |
| 3.3.3 | <i>Support Vector Machine</i> | 28 |
| 3.3.4 | Validação Cruzada | 29 |
| 3.4 | Testes Estatísticos | 31 |

| | | |
|------------|--|-----------|
| 3.4.1 | Teste de <i>Shapiro-Wilk</i> | 31 |
| 3.4.2 | Teste F | 31 |
| 3.4.3 | Teste <i>t-Student</i> | 32 |
| 3.4.4 | Teste das Somas dos Postos de <i>Wilcoxon</i> | 32 |
| 4 | RESULTADOS | 33 |
| 4.1 | Comparativos e Testes estatísticos entre MLP, RBF e SVM | 33 |
| 4.1.1 | Comparação da Raiz do Erro Médio Quadrático | 33 |
| 4.1.2 | Testes Estatísticos | 36 |
| 4.1.2.1 | Arquiteturas com 6 valores de entrada | 36 |
| 4.1.2.2 | Arquiteturas com 12 valores de entrada | 37 |
| 4.1.2.3 | Arquiteturas com 24 valores de entrada | 37 |
| 4.1.2.4 | Arquiteturas com 48 valores de entrada | 38 |
| 5 | CONSIDERAÇÕES FINAIS | 40 |
| 5.1 | Conclusões | 40 |
| 5.2 | Trabalhos Futuros | 40 |
| | REFERÊNCIAS | 42 |

1 Introdução

Este capítulo apresenta a motivação para a realização deste trabalho. Em seguida, são especificados os objetivos a serem atingidos e uma explicação detalhada do conteúdo dos capítulos seguintes.

1.1 Motivação

A expansão acentuada do consumo de energia, embora demonstre o aquecimento econômico e a melhoria da qualidade de vida, tem aspectos negativos. Dentre eles podemos citar: a possibilidade de esgotamento dos recursos utilizados para produção de energia, o impacto ao meio ambiente, e os elevados investimentos em pesquisa e desenvolvimento de novas fontes de geração de energia elétrica (ANEEL, 2008). Atualmente, entre as fontes de energia mais utilizadas estão o petróleo, gás natural e carvão mineral que, além de serem extremamente danosas ao ambiente, são ditas não renováveis, ou seja, seu processo de regeneração é muito lento ou inexistente, tornando sua utilização sustentável inviável.

Desde o início dos anos 90 cientistas e estudiosos alertam para efeitos danosos ao ambiente provocados pela ação humana, tendo como exemplo o aquecimento global, provocado pela emissão elevada de gases causadores do efeito estufa. Estes liberados em larga escala através da queima de combustíveis fósseis para a produção de calor, vapor ou energia elétrica (ANEEL, 2008). Sendo assim, um grande desafio enfrentado pela humanidade é diversificação da matriz energética com fontes renováveis de energia, como energia solar, energia eólica e biomassa.

A utilização do vento como fonte de energia tem seu potencial diretamente ligado à velocidade dele (ANEEL, 2005) e a outras incertezas como direção do vento, clima e topografia. Tais condições justificam a construção de um sistema de previsão de potência eólica, a fim de que se obtenham boas estimativas do potencial eólico em um determinado local, possibilitando a elaboração de boas estratégias para um aproveitamento mais eficiente dessa fonte.

A potência eólica pode ser representada na forma de uma série temporal (ST), que é um conjunto de observações ordenadas no tempo e registradas em intervalos regulares, feitas sobre um evento. Por exemplo, o número de atendimentos diários em um Pronto Socorro ou a quantidade mensal de casos notificados de uma doença específica. Na análise de uma ST, deseja-se modelar o evento estudado para realizar estimativas e descrever o seu comportamento através de testes estatísticos (CARDOSO, 2001). A característica mais importante das STs é a dependência entre as observações vizinhas e a possibilidade de

analisar esta dependência. O estudo deste tipo de dados requer o uso técnicas específicas (EHLERS, 2007), como por exemplo Redes Neurais Artificiais (RNAs). As RNAs são modelos matemáticos que possuem a capacidade de aprender a partir de exemplos, vêm se tornando bastante populares em estudos sobre STs.

A expansão do consumo de energia cria a necessidade de diversificar a matriz energética com fontes renováveis de energia. Uma das mais promissoras é a energia eólica, visto que sua capacidade mundial instalada aumentou 1155% entre 1997 e 2007. Assim, muitos trabalhos de pesquisa vêm sendo realizados nesta área, com o intuito de desenvolver um sistema de previsão eficiente, podemos citar como exemplo (RODRIGUES, 2007), onde utiliza-se a rede *Multi-Layer Perceptron* (MLP). Porém, são implementadas várias configurações com esse tipo de RNA.

A produção de energia elétrica pelas fontes mais utilizadas atualmente, carvão mineral, petróleo e gás, tem consequências negativas, como por exemplo os danos causados ao meio ambiente e fato dessas fontes não serem renováveis. Sendo assim, é necessário introduzir na matriz energética fontes que não apresentem esses pontos negativos, alguns exemplos são: energia eólica, solar, entre outras. A eólica em particular está ligada a uma série de incertezas, tais como velocidade e direção do vento, clima e topografia do local onde será instalado o parque eólico. Com isso, para definir boas estratégias que permitam o uso mais eficiente do potencial eólico do local é necessário ter boas estimativas das potências geradas, sendo fundamental construir um sistema de previsão de potência eólica. Esse sistema deve utilizar métodos que o possibilitem fornecer estimativas precisas sobre a potência eólica.

1.2 Objetivos

1.2.1 Objetivos Gerais

Este trabalho tem como objetivo geral desenvolver um sistema para previsão de potência eólica, utilizando para isso RNAs de arquitetura RBF (LUDERMIR et al., 2007), MLP (VALENÇA, 2009) e SVM (ANDREOLA, 2009), realizar uma comparação de desempenho entre estes métodos que utilizam aprendizagem de máquina e analisar o impacto de diferentes horizontes de entrada (ou seja, quantidade de informações passadas) em cada previsão, para um horizonte de três horas à frente. Pois, a avaliação em variações de curta duração apresenta valores que se aproximam mais da realidade do que a avaliação realizada em variações sazonais e anuais.

1.2.2 Objetivos Específicos

1. Configurar diversas redes.

2. Calcular a correlação entre as variáveis presentes na base de dados.
3. Encontrar os valores para os parâmetros que apresentem a previsão mais eficiente, para um intervalo de três horas à frente.
4. Comparar os resultados obtidos pelas três arquiteturas de rede utilizando testes estatísticos.

1.3 Estrutura da Monografia

O capítulo 2 apresenta os conceitos necessários para o entendimento dos experimentos realizados: energia eólica e Redes Neurais Artificiais, enfatizando as três arquiteturas escolhidas para esse trabalho, a MLP, a RBF e a SVM. No capítulo 3 é descrita a metodologia escolhida com o intuito de alcançar os objetivos definidos na seção 1.2. Além de explicar como se dará o processamento dos dados, as configurações e uso das três arquiteturas selecionadas para serem utilizadas no trabalho. O capítulo 4 apresenta os resultados atingidos após as simulações, fazendo uma comparação entre os resultados alcançados com as três arquiteturas. E finalmente o capítulo 5 mostra as conclusões inferidas a partir dos resultados obtidos e propostas de trabalhos futuros.

2 Fundamentação Teórica

Este capítulo fornece uma breve explicação dos conceitos chaves para o completo entendimento dos experimentos propostos. A seção 2.1 discorre sobre a energia eólica, apresentando desde sua demanda no Brasil e no mundo até informações sobre sua geração, equipamentos utilizados para produção de energia elétrica e considerações sobre o vento. A seção 2.2 traz informações sobre as Redes Neurais Artificiais (RNAs), explicando as arquiteturas escolhidas para serem utilizadas neste trabalho (MLP, RBF e SVM) e o modo como é realizado o treinamento em cada uma delas.

2.1 Energia Eólica

A energia eólica tem sua geração diretamente ligada ao vento, resultando do deslocamento de massas de ar (FONSECA et al., 2012). A ideia principal é aproveitar a energia cinética resultante desses deslocamentos. Esses causados pela diferença de temperatura na superfície do planeta (ANEEL, 2008), ocasionada pela energia solar: enquanto o sol aquece a água e a terra de um lado do planeta, o outro lado é resfriado. Porém, regiões diferentes respondem de maneiras diferentes a esse aquecimento. Por exemplo, os oceanos se aquecerão mais lentamente do que as terras porque a água tem uma capacidade maior de estocar calor, gerando massas de ar com diferentes características. Devido as causas que lhe dão origem a energia eólica é considerada renovável.

2.1.1 Histórico

O aproveitamento dos benefícios da energia eólica pelo homem data da antiguidade. Acredita-se que foram os egípcios os primeiros a utilizarem o vento. Barcos encontrados em um túmulo da época do ano 4000 a.c. indicam que nessa época já se fazia o uso de energia eólica em embarcações (FONSECA et al., 2012). Mas estas ainda utilizavam os remos em conjunto com o vento para se locomover. Os primeiros a utilizarem barcos movidos exclusivamente com energia eólica foram os fenícios por volta do ano 1000 a.c..

Contudo, na antiguidade o uso desta forma de energia não se resumia a sua conversão em energia cinética, mas contempla também a sua conversão em energia mecânica através dos moinhos de ventos, utilizados há mais de 1000 anos, e tem sua origem ligada ao Oriente Médio por volta do século VII. Alguns indícios datam o uso do moinho de vento do século X na Pérsia (Figura 1). Posteriormente, foram trazidos para a Europa pelas Cruzadas na Idade Média, onde incorporavam velas e eixo horizontal, designadas *Sail Windmill* (ver Figura 2).



Figura 1 – Moinho de vento Persa

[Fonte: Reproduzido de (HECKL, 2015)]



Figura 2 – Sail Windmill

[Fonte: Reproduzido de (HECKL, 2015)]

Após algumas evoluções surgiram os *Post mills* (ver Figura 3) e os *Dutch windmills*, usados essencialmente para bombear água e cultivar a terra. A partir do século XIX nos Estados Unidos, passou-se a utilizar turbinas eólicas multi-lâmina, mostrada na Figura 4, para a irrigação.



Figura 3 – Post mill

[Fonte: Reproduzido de (FERREIRA, 2011)]



Figura 4 – Turbina eólica americana

[Fonte: Reproduzido de (BOLDUAN, 2000)]

A Charles F. Brush é atribuída a invenção do primeiro aerogerador automático. Em 1888, ele construiu a turbina eólica de Brush, hoje reconhecida como a primeira turbina de vento de funcionamento automatizado para produção elétrica. Possuía um rotor com um diâmetro de 17 metros (o maior do mundo na altura) e 177 lâminas feitas de madeira de cedro, mostrada na Figura 5. Apesar do seu tamanho, a máquina tinha uma potência de 12 KW, não conseguindo uma grande eficiência. Foi o dinamarquês Poul La Cour a conseguir desenvolver as turbinas de rotação rápida, em 1897, construiu os seus próprios túneis de vento para as suas experiências. Utilizou a energia elétrica proveniente das turbinas

eólicas para produzir hidrogênio através da eletrólise. Esses experimentos permitiram o desenvolvimento das turbinas de geração de energia eólica ao longo do século XX.

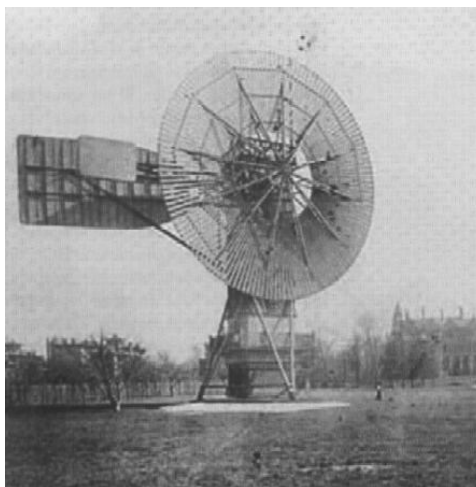


Figura 5 – Turbina eólica de Charles Brush

[Fonte: Reproduzido de (MOLENAAR, 2003)]

2.1.2 Geração de Energia Eólica

A energia eólica tem sua origem na energia solar, uma vez que os ventos são os movimentos das massas de ar resultantes do aquecimento não uniforme da superfície terrestre (FERREIRA, 2011). As regiões tropicais são mais aquecidas que as regiões polares, pois estão mais próximas do sol. Assim, o ar quente que se encontra nas baixas altitudes das regiões tropicais tende a subir, sendo substituído por uma massa de ar mais frio que se desloca das regiões polares, Figura 6. Esses deslocamentos geram os ventos do planeta.

Os ventos mais fortes e constantes são encontrados a mais de um quilômetro da superfície terrestre. No entanto, só é possível instalar aproveitamentos eólicos a algumas dezenas de metros da superfície. A densidade do ar, a intensidade, direção e velocidade do vento relacionam-se com aspectos geográficos naturais como relevo vegetação e interações térmicas entre a superfície da terra e a atmosfera. Sendo assim, a obtenção de energia eólica, a exemplo do que ocorre com outras fontes como a hidráulica, exige a existência de condições naturais favoráveis e específicas. A avaliação destas condições requer trabalhos sistemáticos de coleta e análise de dados sobre o regime de vento do local (FERREIRA, 2011) (ANEEL, 2008).

A geração eólica ocorre pelo contato do vento com as pás do cata-vento, elementos integrantes da usina. Ao girar, essas pás dão origem à energia mecânica que aciona o rotor do aerogerador, produzindo eletricidade. A quantidade de energia mecânica transferida e o



Figura 6 – Movimento das massas de ar

[Fonte: Reproduzido de
<<http://www.coladaweb.com/geografia/pressao-atmosferica-massas-de-ar-precipitacao-e-ventos>>]

potencial de energia elétrica a ser produzida estão diretamente relacionados à densidade do ar, à área coberta pela rotação das pás e à velocidade do vento.

A evolução tecnológica permitiu o desenvolvimento de equipamentos mais potentes, que aproveitam o vento de maneira mais eficiente. Essa evolução vai desde o aumento do diâmetro das turbinas até o aumento da altura das torres, que inicialmente era de 10 metros aproximadamente e hoje supera os 50 metros, incluindo também sistemas avançados de transmissão, melhor aerodinâmica, estratégias de controle e operação das turbinas. Em 1985 a altura das torres era de 20 metros, produzindo uma potência média de 50 kW (quilowatts) e hoje chega a superar os 100 metros, permitindo a obtenção em uma única turbina de 5000 kW.

Para um melhor aproveitamento da energia eólica é importante distinguir os tipos de variações temporais da velocidade dos ventos, pois estes a influenciam fortemente. Essas variações podem ser: variações anuais, sazonais, diárias e de curta duração (RODRIGUES, 2007). A seguir uma breve descrição de cada uma:

- **Variações Anuais** - Para se obter um bom conhecimento do regime dos ventos não é suficiente basear-se na análise de dados de vento de apenas um ano; o ideal é dispor de dados referentes a vários anos.
- **Variações Sazonais** - O aquecimento não uniforme da superfície terrestre resulta em significativas variações no regime dos ventos, resultando na existência de diferentes estações do ano. Sendo assim, a utilização de médias anuais (ao invés de médias sazonais) pode levar a resultados que se afastam da realidade.

- **Variações Diárias** - Essas variações são importantes para definir o local mais adequado para a instalação do sistema eólico dentro de uma área que já foi escolhida para hospedar um parque eólico. Ao comparar a evolução da velocidade média ao longo do dia percebe-se que há uma significativa variação de um mês para os outros. Com esse tipo de informação pode-se projetar melhor o sistema eólico.
- **Variações de Curta Duração** - As variações de curta duração estão associadas tanto às pequenas flutuações quanto às rajadas de vento. Em um primeiro momento, essas variações não são consideradas na análise do potencial eólico de uma região.

Além da velocidade, outro parâmetro que influencia muito a produção de energia eólica é a direção do vento. Frequentes mudanças na direção indicam a existência de rajadas de ventos, e a medida da direção do vento auxilia a localização das turbinas nos parque eólicos.

2.1.3 Turbina Eólica

Como pode ser visto na seção 2.1.1, os egípcios foram os primeiros a utilizar a energia eólica. Eles a utilizaram por volta de 2800 a.c. para mover as embarcações em conjunto com a força dos remos. Com o passar dos anos as técnicas de aproveitamento da energia eólicas foram evoluindo até que, no do início do século XX, essa energia começou a ser utilizada para produzir eletricidade, sendo as primeiras tentativas creditadas aos dinamarqueses. Por volta da década de 1930, cerca de um dúzia de firmas começou a produzir e vender esses geradores primitivos, que geravam em torno de 1kW.

Muitos países europeus começaram a construir avançados projetos de geradores eólicos entre 1950 e 1960. A primeira turbina eólica comercial ligada a rede elétrica pública foi instalada em 1976, na Dinamarca. Uma das mais memoráveis foi construída perto de Rutland, Vermont, USA e foi projetada para fornecer 1250 kW para Vermont.

As turbinas eólicas são máquinas que utilizam a energia eólica - energia cinética dos ventos - para rotacionar sua pás, transformando essa energia em energia mecânica ou elétrica. São classificadas de acordo com a posição do eixo do rotor (TABLADA, 2010) em turbinas eólicas de eixo vertical, possuem pás que giram em um plano paralelo ao do vento, e turbinas eólicas de eixo horizontal, possuem pás que giram em um plano perpendicular ao do vento. Elas são compostas pelo rotor e pela torre que o sustenta, pela transmissão e caixa multiplicadora - responsável por transmitir a potência fornecida pelo rotor ao gerador - e pelo conversor (FONSECA et al., 2012).

A energia cinética só é extraída do ar que passa pela área interceptada pelas pás rotativas. A energia cinética bruta por unidade de tempo do vento passando por uma área A perpendicular ao seu vetor velocidade instantânea V , é dada pela equação 2.1:

$$P = C_p \cdot \frac{1}{2} \cdot \rho \cdot A \cdot V^3 \quad (2.1)$$

Temos:

- ρ = densidade do ar em kg/m^3 que varia de acordo com a latitude e as condições atmosféricas;
- C_p = é o coeficiente de desempenho que se relaciona com a energia cinética de saída e depende do modelo e da relação entre a velocidade do rotor e a velocidade do vento;
- V = velocidade do vento em m/s;
- A = área em m^2 .

2.1.3.1 Turbinas Eólicas de Eixo Vertical

Esse tipo de turbina foi inventado pelo engenheiro francês Darrieus, incluindo duas lâminas, como pode ser observado na Figura 7. Esta turbina é unidirecional, ou seja, aceita o vento de qualquer direção, diferente do que acontece com as turbinas convencionais que se ajustam a direção do vento. Por isso, as turbinas de eixo vertical não necessitam de sistemas de controle de direção, o que simplifica bastante os mecanismos de transmissão (RODRIGUES, 2007). Algumas grandes vantagens dessas turbinas em relação a outros tipos são a manutenção, muito mais prática, e uma variabilidade de aplicações elétricas e mecânicas muito maior. Isso se deve ao fato de o rotor e suas partes elétricas se localizarem na parte inferior da turbina.

2.1.3.2 Turbinas Eólicas de Eixo Horizontal

A origem da turbina de eixo horizontal é atribuída a Rússia no ano de 1931, sendo construída junto ao Mar Negro. Essa turbina tinha um rotor de três pás, como pode ser observado na Figura 8, com um diâmetro de 30 metros e uma potência nominal de 100 kW. Funcionou durante cerca de dois anos ligada a rede de corrente alternada. Em sistema para aproveitamento com ventos de baixa velocidade é comum utilizar-se o rotor de múltiplas pás, ver Figura 9.

2.1.4 Energia Eólica no Mundo

O interesse pelo uso da energia eólica vem crescendo consideravelmente, principalmente na Europa. Trinta por cento de toda a capacidade nova anual instalada na União Europeia, entre 2002 e 2006, foi de geração de energia eólica (SALLES, 2009). Em 2007



Figura 7 – Turbina de Darrieus de duas lâminas

[Fonte: Reproduzido de (FERREIRA, 2011)]



Figura 8 – Turbina eólica de eixo horizontal

[Fonte: Reproduzido de (RODRIGUES, 2007)]



Figura 9 – Aerogerador com múltiplas pás

[Fonte: Reproduzido de
<<http://www.engquimicasantosp.com.br/2013/12/energia-eolica-e-aerogeradores.html>>]

essa participação anual subiu para 40% da capacidade total instalada, tornando-a a forma de geração de energia elétrica que mais cresce na Europa.

Atualmente existem mais de 30 mil turbinas eólicas em operação no mundo (VITERBO, 2008). A Associação Europeia de Energia Eólica estabeleceu, em 1991, como metas a instalação de 4.000 MW (Megawatts) de energia eólica na Europa até o ano 2000 e 11.500 MW até 2005. Essas metas foram atingidas muito antes do esperado, 4000 MW em 1996 e 11500 MW em 2001. O mercado vem crescendo substancialmente, principalmente na Alemanha, EUA, Dinamarca e Espanha, onde a potência adicionada anualmente supera 3.000 MW. Esse crescimento de mercado fez com que a Associação Europeia de Energia Eólica estabelecesse uma nova meta que é em 2020 12% da energia do mundo ser gerada pelo vento, com uma capacidade instalada de 1200 GW (Gigawatts) (ANEEL, 2005). O Conselho Europeu de Energia Renovável (EREC) traçou uma meta para as fontes renováveis. Essa meta é que em 2022 a fonte eólica possa atingir um nível de 4000 TWh ao ano, superando a fonte hidráulica de porte como a principal fonte renovável de energia elétrica do mundo.

Na América do Norte a instalação de fazendas eólicas se intensificou apenas por volta de 2003, sendo que no final de 2008 os Estados Unidos se tornaram o país com maior capacidade de energia eólica instalada de geração de energia eólica. O interesse pela instalação de novas fazendas eólica continua crescendo. Somente na costa da Alemanha planeja-se instalar gradativamente 25 GW até 2020, porém, devido as novas diretrizes do

Governo alemão, este valor passou por 10 GW.

2.1.5 Energia Eólica no Brasil

No Brasil, as primeiras unidades utilizadas para a medição de energia eólica foram instaladas no início dos anos 1990, no Ceará e em Fernando de Noronha. Essas medições determinaram o potencial eólico local e a instalação das primeiras turbinas eólicas no Brasil. O primeiro estudo sobre a energia eólica no Brasil foi a elaboração do Atlas Eólico do Nordeste, publicado pelo Centro Brasileiro de Energia Eólica (CBEE) com o apoio da Agência Nacional de Energia Elétrica (ANEEL), em 1998. Apesar das divergências entre especialistas e instituições, as estimativas sobre o potencial eólico brasileiro apresenta valores extremamente relevantes. As estimativas, que até poucos anos atrás eram de 20000 MW, hoje indicam valores superiores a 60000 MW. A falta de informação e a diferença nas metodologias aplicadas em cada estudo provoca essas divergências. Para ampliar os conhecimentos sobre o potencial eólico brasileiro, vários estudos vem sendo realizados, como, por exemplo atlas eólicos para alguns estados.

A participação da energia eólica na geração de energia elétrica do país ainda é pequena. Em setembro de 2003 havia apenas 6 centrais eólicas em operação no país, perfazendo uma capacidade instalada de 22.075 kW. Entre essas centrais, destacam-se Taíba e Prainha, no Estado do Ceará, que representam 68% do parque eólico nacional.

Porém, existem alguns incentivos vigentes para o setor elétrico que devem despertar um maior interesse por investimentos e projetos na área. Entre esses incentivos podemos destacar o Programa de Incentivo às Fontes Alternativas de Energia Elétrica (PROINFA) (MME, 2004), cujo objetivo é aumentar a participação da energia elétrica produzida por empreendimentos concebidos com base em fontes eólica, biomassa e pequenas centrais hidrelétricas no Sistema Elétrico Interligado Nacional (SIN). Outro incentivo importante é a possibilidade de complementaridade entre a geração hidrelétrica e a eólica, visto que o maior potencial eólico, na região Nordeste, ocorre durante o período de estiagem, conforme ilustrado na Figura 10.

O Brasil é favorecido em termos de ventos, que se caracterizam por uma presença duas vezes superior à média mundial e pela volatilidade de 5% (oscilação da velocidade), o que dá maior previsibilidade ao volume a ser produzido. Além disso, como dito no parágrafo anterior, a velocidade costuma ser maior em períodos de estiagem, sendo assim, é possível operar as usinas eólicas em sistema complementar com as usinas hidrelétricas, de forma a preservar a água dos reservatórios em períodos de poucas chuvas. Finalmente, estimativas mostradas em (AMARANTE et al., 2001) apontam para um potencial de geração de energia eólica de 143 mil MW no Brasil, como pode ser visto na Figura 11.

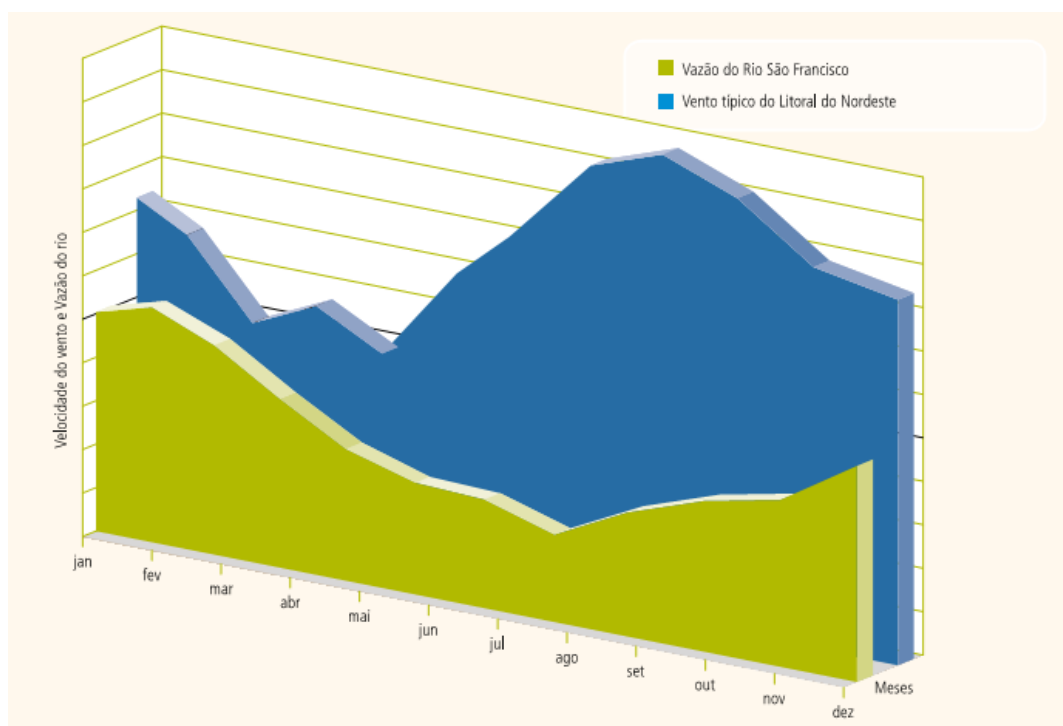


Figura 10 – Complementaridade entre geração hidrelétrica e eólica

[Fonte: Reproduzido de (ANEEL, 2005)]



Figura 11 – Potencial eólico brasileiro

[Fonte: Reproduzido de (ANEEL, 2008)]

2.2 Redes Neurais

2.2.1 Introdução

Redes neurais artificiais são modelos matemáticos - inspirados no sistema nervoso dos seres vivos - que tem a capacidade de aprender através de exemplos. Possuem a capacidade de aquisição e manutenção do conhecimento (SILVA et al., 2010). Quanto à sua constituição, pode-se definir uma RNA como um sistema composto por elementos de processamento interconectados chamados de neurônios, dispostos em camadas (VALENÇA, 2009).

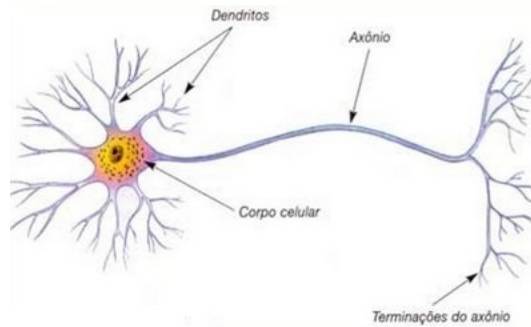
2.2.1.1 Histórico

A primeira publicação relacionada à redes neurais data de 1943, através de um artigo elaborado por McCulloch & Pitts (VALENÇA, 2009). Esse trabalho resultou na construção do primeiro neurônio artificial, resultado da modelagem matemática inspirada no neurônio biológico. A partir daí outros pesquisadores continuaram o desenvolvimento de modelos matemáticos baseados no neurônio biológico, dando origem a diversas arquiteturas (SILVA et al., 2010).

Entre as primeiras arquiteturas que surgiram pode-se destacar a pesquisa de Frank Rosenblatt que, entre 1957 e 1958, desenvolveu o primeiro neuro-computador, denominado *Mark I - Perceptron*, idealizando o modelo básico do *Perceptron*, capaz de reconhecer padrões simples. Em 1960, Widrow e Hoff criaram uma regra de aprendizagem para uma extensão do *Perceptron* chamada de ADALINE. Mas, em 1969, verificou-se que o *Perceptron* conseguia distinguir apenas padrões linearmente separáveis. Este fato gerou um desinteresse pelas RNAs, que ressurgiram graças ao trabalho publicado por John Hopfield em 1982 e ao desenvolvimento do algoritmo *backpropagation* (VALENÇA, 2009).

2.2.1.2 Neurônio Biológico

A célula elementar do sistema nervoso cerebral é o neurônio, seu papel pode ser resumido em conduzir impulsos. Constitui-se de três partes: dendritos, corpo celular e no axônio, ver Figura 12. A função dos dendritos é captar os estímulos vindos de outros neurônios ou do próprio meio onde o neurônio se encontra. O corpo celular, por sua vez, tem a função de processar todas as informações vindas dos dendritos a fim de produzir um potencial de ativação. Este potencial indicará se o neurônio poderá disparar um impulso elétrico ao longo de seu axônio. Finalmente, o axônio tem a missão de conduzir os impulsos elétricos para outros neurônios através das sua terminação que contém ramificações denominadas terminações sinápticas.

**Figura 12 – Neurônio Biológico**

[Fonte: Reproduzido de
 <<http://www.biologia.seed.pr.gov.br/modules/galeria/detalhe.php?foto=256&evento=3>>]

2.2.1.3 Neurônio Artificial

Os elementos computacionais ou unidades processadoras, denominadas neurônios artificiais, são modelos simplificados dos neurônios biológicos (SILVA et al., 2010). Os neurônios artificiais são não-lineares, fornecem saídas contínuas, e exercem funções simples, como coletar os sinais existentes em suas entradas, agregá-los de acordo com sua função operacional e produzir uma resposta.

A representação do neurônio artificial pode ser vista na Figura 13. Nesta representação podemos ver as entradas $(x_1, x_2, x_3, \dots, x_m)$, representando os sinais recebidos pelo neurônio do meio externo ou de outros neurônios. Temos ainda os pesos $(w_1, w_2, w_3$ e $w_m)$, que representam o grau de importância de cada entrada para o neurônio, ou seja, quanto maior o peso mais importante é a entrada para o neurônio. O somatório de todos os produtos entrada x peso, menos o *bias*, gera um resultado v_i , como pode ser visto na equação 2.2. Esse resultado será o parâmetro para a função de ativação $\varphi(\cdot)$, que gera a saída do neurônio, (VALENÇA, 2009).

$$v_i = \sum_{j=1}^n w_{ij} \cdot x_j - bias \quad (2.2)$$

2.2.2 Multi-Layer Perceptron

Multi-Layer Perceptron (MLP) é uma RNA composta por nós agrupados em camadas: uma camada de entrada, uma ou mais camadas intermediárias e uma camada de saída. A Figura 14 mostra uma MLP com duas camadas escondidas. MLP é uma rede onde cada nó de uma camada está conectado unicamente aos nós da camada seguinte, sem realimentação. Assim, o sinal se propaga pela rede camada a camada, a partir da camada de entrada até a camada de saída.

As conexões entre os nós são conhecidas como sinapses. A camada de entrada é a

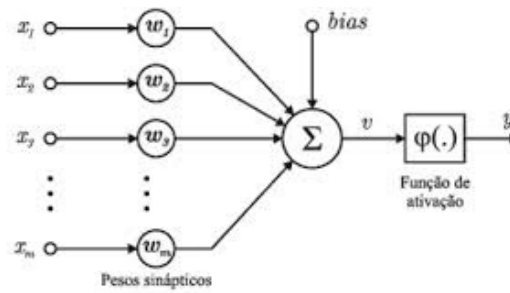


Figura 13 – Neurônio Artificial

[Fonte: Reproduzido de <<http://www.scielo.br/img/revistas/ca/v21n4/a03fig17.gif>>]

que trata a alimentação das características, ou seja, recebe o sinal de entrada, e o número de nós nessa camada depende da dimensionalidade do espaço de observação (LEITE, 2012). A camada intermediária é a responsável pela não linearidade da rede e que permite que as redes MLP sejam capazes de resolver problemas reais. Por sua vez, a camada de saída é responsável pela resposta da rede e representa a variável desejada.

As redes MLP são ainda caracterizadas pelas elevadas possibilidades de aplicações em diversos tipos de problemas relacionados com as mais diferentes áreas do conhecimento, sendo também consideradas uma das arquiteturas mais versáteis quanto aplicabilidade (SILVA et al., 2010):

- Aproximação universal de funções.
- Reconhecimento de padrões.
- Identificação e controle de processos.
- Previsão de séries temporais.
- Otimização de sistemas.

As redes MLP são redes de treinamento supervisionado. Para o seu treinamento é necessário utilizar um algoritmo que permita estabelecer um conjunto de pesos ótimos para a rede. O que torna o treinamento da MLP complexo é a existência de pelo menos uma camada intermediária, pois, com a inclusão dessas camadas, não se conhece o erro necessário para realizar o reajuste dos pesos. Sendo assim, para o treinamento dessas redes é utilizado o *backpropagation*, que resolve este problema realizando uma propagação recursiva dos erros.

De acordo com Valença (2009), o *backpropagation* é realizado em dois passos: no primeiro passo os sinais são propagados da camada de entrada para a camada de saída (sentido progressivo), calculando-se o sinal de saída e o erro, sem alterar os pesos; no

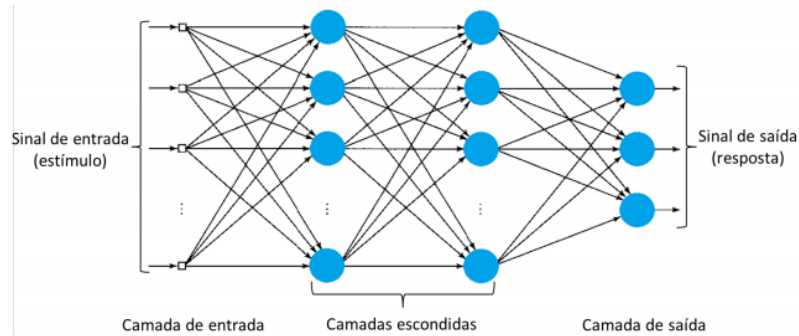


Figura 14 – Arquitetura da rede MLP

[Fonte: Reproduzido de Leite (2012)]

segundo passo, os erros são propagados da camada de saída para a entrada (recursivamente) e os pesos são ajustados.

Algoritmo Backpropagation

O algoritmo *backpropagation* é um algoritmo supervisionado que utiliza pares (entrada, saída desejada) para ajustar os pesos da rede. O treinamento ocorre em duas fases, onde em cada fase percorre a rede em um sentido. Estas fases são chamadas de *forward* e *backward*. A primeira fase, *forward*, é utilizada para calcular a saída da rede, enquanto a segunda fase, *backward*, atualiza os pesos das conexões da rede (LUDERMIR et al., 2007). A Figura 15 ilustra o sentido em que as duas fases atuam na rede.

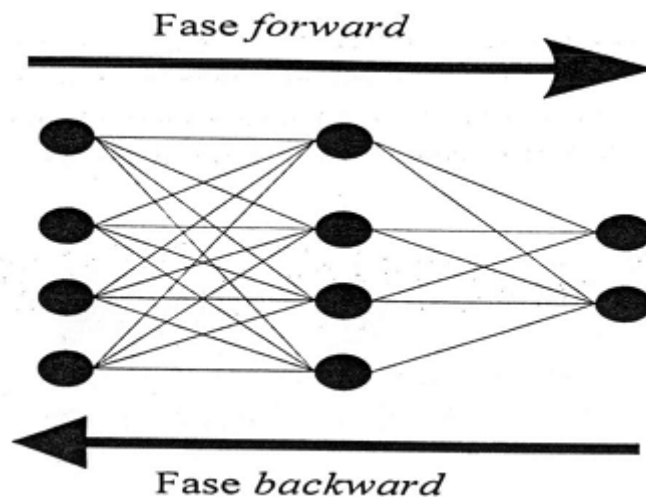


Figura 15 – Fases do algoritmo *Backpropagation*

[Fonte: Reproduzido de Ludermir et al. (2007)]

Fase *Forward*

A primeira fase do algoritmo corresponde à propagação do sinal da camada de entrada para a camada de saída e com o cálculo do erro na saída, que será utilizado para

ajustar os pesos. A equação 2.3 mostra como é calculado o valor do sinal de cada neurônio na camada de saída:

$$y_i^2 = f^2 \cdot (net_i^2) \quad (2.3)$$

Temos:

- y_i^2 - representa o valor do sinal do neurônio i da camada 2, camada de saída para uma rede com uma camada escondida;
- f^2 - função de ativação da camada 2;
- net_i^2 - valor do sinal recebido pelo neurônio i da camada 2.

A equação 2.4 mostra como é calculado o erro para cada neurônio na saída:

$$e_i(n) = (d_i - y_i^2) \quad (2.4)$$

Onde $e_i(n)$ é o valor do erro do neurônio i da camada de saída, para o exemplo n ; d_i é o valor desejado para o neurônio i e y_i^2 é o valor calculado no neurônio i da camada 2.

Fase *Backward*

Nesta fase os valores dos pesos da rede são ajustados levando em consideração o erro calculado na fase anterior. Esse ajuste nos pesos é realizado de acordo com a equação 2.5:

$$w_{ij}^m(novo) = w_{ij}^m(antigo) + \alpha \delta_i^m f^{m-1}(net_j^{m-1}) \quad (2.5)$$

Onde w_{ij}^m representa os pesos das conexões entre os neurônios, com i representando o neurônio que emite o sinal e j o neurônio que recebe o sinal; α é a taxa de aprendizagem e δ_i^m é a sensibilidade do neurônio.

Para o cálculo da sensibilidade é preciso levar em conta duas situações: quando o neurônio de índice j está na camada de saída, equação 2.6, e quando não está, equação 2.7.

$$\delta_i^2 = f^{2'}(net_i^2) \cdot e_i(n) \quad (2.6)$$

$$\delta_j^{m-1} = f^{m-1'}(net_j^{m-1}) \sum_{i=1}^N w_{ij}^m \delta_i^m \quad (2.7)$$

Através da equação 2.7 pode-se concluir que o valor da sensibilidade é calculado recursivamente da última para a primeira camada. Nesta equação $f^{m-1'}(net_j^{m-1})$ representa

a derivada da função de ativação da camada que emite o sinal, o índice i representa o número do neurônios da camada que recebe o sinal, ou seja, da camada que está imediatamente a frente e tem N neurônios e sensibilidade δ_i^m (PAZ, 2014).

2.2.3 Radial Basis Function

As redes conhecidas como RBF (*Radial Basis Function*), podem ser empregadas em quase todos os tipos de problemas tratados pela MLP, inclusive aproximação de funções e reconhecimento de padrões (SILVA et al., 2010). Mas, diferentemente da MLP que pode ser composta de diversas camadas intermediárias, a estrutura típica de uma RBF possui apenas uma, cujas funções de ativação são do tipo gaussiana, como pode ser observado na Figura 16. De acordo com Ludermit et al. (2007), a utilização deste tipo função - de base radial - na camada intermediária é que dá o nome à rede, enquanto que na rede MLP utiliza-se na camada intermediária, geralmente a função sigmoideal.

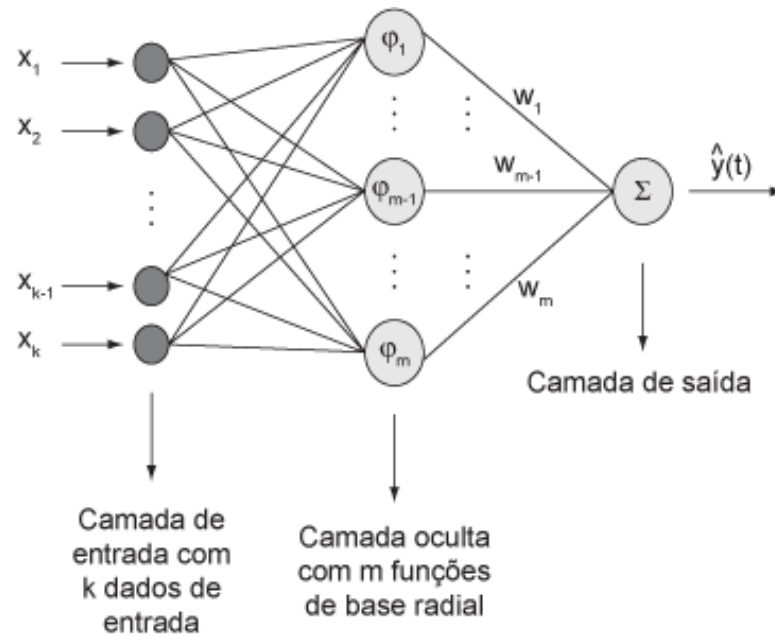


Figura 16 – Estrutura da rede RBF

[Fonte: Reproduzido de Coelho, Santos e Jr (2008)]

Cada camada da rede RBF tem um papel específico para o correto comportamento da rede. A primeira camada representa as entradas (os sinais) recebidas pela rede. A segunda camada, cujos neurônios tem função de base radial, agrupa os dados de entrada em *clusters*, transformando o conjunto de entradas não linearmente separáveis em um conjunto de saídas linearmente separáveis. A camada de saída procura classificar os padrões recebidos da camada anterior.

Funções de base radial

Segundo Aliaga (2010), funções de base radial são uma classe de funções em que a sua resposta decresce, ou cresce, monotonicamente com o distanciamento de um ponto central. As funções mais utilizadas são:

- Funções Gaussianas:

$$\phi_i(x) = \exp\left(-\frac{\|x_j - \mu_i\|^2}{2\sigma_i^2}\right) \quad (2.8)$$

- Funções Multi-Quadráticas:

$$\phi_i(x) = \sqrt{\|x_j - \mu_i\|^2 + 2\sigma_i^2} \quad (2.9)$$

- Funções Multi-Quadráticas Inversas:

$$\phi_i(x) = \frac{1}{\sqrt{\|x_j - \mu_i\|^2 + 2\sigma_i^2}} \quad (2.10)$$

- Função de base Lâmina *Spline* fina:

$$\phi_i(x) = \frac{\|x_j - \mu_i\|}{\sigma_i^2} \log \frac{\|x_j - \mu_i\|}{\sigma_i} \quad (2.11)$$

Temos: x_j são as entrada da rede e μ_i e σ_i são respectivamente o centro e a largura da i -ésima função.

Treinamento da rede RBF

O treinamento na rede RBF é efetivado de forma supervisionada e é dividido em dois estágios distintos: o primeiro é associado aos ajustes dos pesos dos neurônios da camada intermediária e depende apenas das características dos dados de entrada; já o segundo é associado ao ajuste dos pesos dos neurônios da camada de saída. Outra característica importante do treinamento dessas redes é que esse tem início na camada intermediária e se encerra na camada de saída (SILVA et al., 2010).

Primeira fase

Esta etapa geralmente é realizada de forma não supervisionada, e como dito anteriormente tem efeito sobre a camada escondida. O objetivo é ajustar os parâmetros da função radial (ALIAGA, 2010). Os parâmetros a serem ajustados são:

- Centros (μ_i): existem duas forma de escolher os centros. A primeira forma é distribuir de forma randômica, e a segunda forma é aplicar algum algoritmo de pré-processamento, como por exemplo Mínimos Quadrados Ortogonais e o *K-means*.
- Largura (σ_i): geralmente a escolha deste valor é realizada de forma empírica e com ajuda de heurísticas, como por exemplo atribuir para cada unidade um valor de largura diferente e cálculo da média do seu centro até o centro da unidade mais próxima.

Segunda fase

Como a camada escondida transforma o conjunto de entradas não linearmente separáveis em um conjunto de saídas linearmente separáveis, o ajuste dos pesos da camada de saída reduz-se a um problema linear. De acordo com (PAZ, 2014), Nesta segunda etapa do treinamento, os pesos que ligam a camada escondida à camada de saída da rede são ajustados através de um combinador linear que atua na camada de saída para que as ativações das funções de base radial gerem uma saída linear, aos moldes dos dados de entrada.

2.2.4 Support Vector Machine

A SVM é uma técnica de aprendizagem de máquina que se baseia no aprendizado estatístico. Esta técnica busca minimizar o erro com relação ao conjunto de treinamento (Risco Empírico), juntamente com o erro com relação ao conjunto de teste. A motivação para esse princípio surgiu da necessidade de desenvolver limites teóricos para a capacidade de generalização dos sistemas de aprendizagem (ANDREOLA, 2009). O objetivo da SVM consiste em encontrar um hiperplano ótimo que possua a maior margem de diferenciação entre as classes, a fim de obter um equilíbrio entre ambos os erros citados anteriormente, e minimizando o excesso de ajustes, melhorando a capacidade de generalização. Inicialmente foi desenvolvida como um método de separação linear, mas é possível estendê-la de forma a separar classes que são não lineares, a ideia principal é transportar os dados para um espaço de dimensão maior no qual eles possam ser separados linearmente (LEITE, 2012).

A função de decisão que maximiza a habilidade de generalização é determinada pelo problema de duas classes, assumindo que as amostras de treinamento das diferentes classes são linearmente separáveis. A função de decisão mais adequada é aquela para a qual a distância entre os conjuntos das amostras de treinamento é maximizada, como pode ser visto na Figura 17. A função de decisão que maximiza esta separação é denominada de ótima.

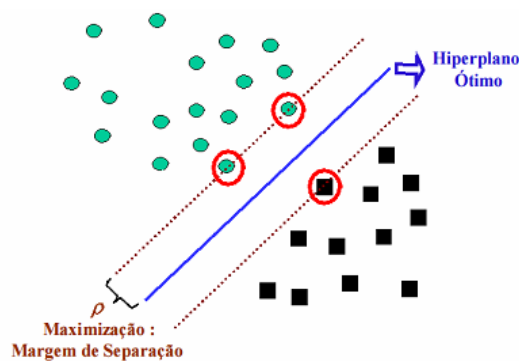


Figura 17 – Hiperplano ótimo separando os dados com a máxima margem ρ

[Fonte: Reproduzido de Andreola (2009)]

Para tratar dos casos não-linearmente separáveis, se introduz a variável de folga (*slack variable*) ξ_i (≥ 0), ver Figura 18. As amostras de treinamento x_I , para as quais $0 < \xi_i < 1$, são corretamente classificadas, embora sem a margem de separação máxima. Por outro lado, as amostras de treinamento x_I , para as quais $\xi_i \geq 1$, são erroneamente classificadas pelo hiperplano ótimo.

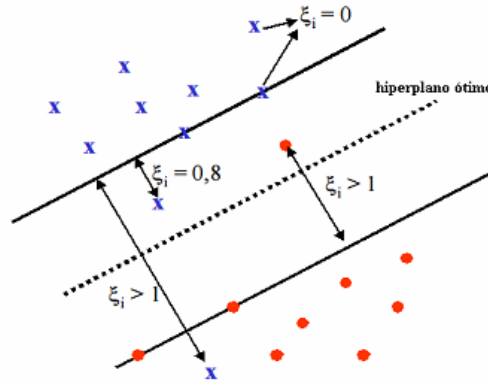


Figura 18 – Exemplos de valores e situações da variável de folga ξ .

[Fonte: Reproduzido de Andreola (2009)]

Outra característica presente na SVM quando trata problemas não linearmente separáveis é a Função de *Kernel*. As funções de *kernel* têm a finalidade de projetar os vetores de características de entrada em um espaço de características de alta dimensão para classificação de problemas que se encontram em espaços não linearmente separáveis (JUNIOR, 2010).

A Figura 19 mostra o processo de transformação de um domínio não linearmente separável, em um problema linearmente separável através do aumento da dimensão, onde é feito um mapeamento por uma função de *kernel* $F(x)$.

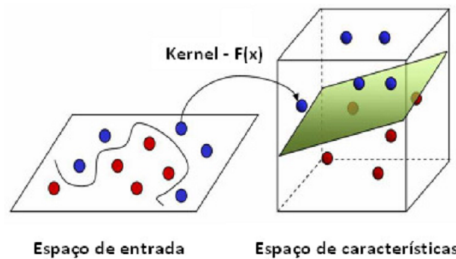


Figura 19 – Transformação de um problema não linearmente separável em um problema linearmente separável

[Fonte: Reproduzido de Junior (2010)]

De acordo com Rodrigues et al. (2007) uma das grandes vantagens da SVM é seu alto poder de generalização. Isto ocorre pois a complexidade da hipótese não depende do

número de atributos, mas sim da margem com que eles separam os dados.

Teoria da Aprendizagem

A Teoria da Aprendizagem Estatística no contexto de SVM, tem o objetivo de controlar a habilidade de generalização da técnica. Existem diversas teorias de aprendizagem que podem ser aplicadas. A técnica SVM é derivada dos princípios básicos descritos na teoria de Vapnik e Chervonenkis (VC). A teoria VC é a mais apropriada para descrever SVM.

Segundo Santos (2002), para uma tarefa de aprendizagem com uma quantidade de dados de treinamento finita, o melhor desempenho de generalização ocorre quando é atingido um equilíbrio entre a precisão alcançada em um conjunto de treinamento particular e a capacidade do sistema.

3 Metodologia

Este capítulo apresenta os passos que foram executados para atingir os objetivos definidos no capítulo 1, e como será aplicada a teoria apresentada no capítulo anterior para a realização do estudo proposto neste trabalho.

3.1 Base de Dados

A base de dados utilizada no desenvolvimento deste trabalho correspondem às medições feitas em um parque eólico. Por se tratarem de informações importantes, o nome do parque eólico, de onde foram coletados os valores das potências, será omitido e um codinome será utilizado.

Os valores das potências eólicas presentes na base de dados foram medidos em intervalos de 30 minutos pra cada 24 horas, totalizando num total de 48 valores de potências por dia. Como foi definido na seção 1.2, este trabalho busca uma previsão em intervalos de trinta minutos, três horas à frente. Tem-se para cada dia além dos 48 valores de potências medidos ao longo do dia, mais 6 valores correspondentes às três horas da previsão realizada. Assim, na base de dados do Parque I, para cada dia tem-se 54 valores de potência (48 entradas e 6 saídas). A base de dados contém observações feitas durante 6966 dias.

3.2 Processamento dos Dados

3.2.1 Normalização dos Dados

Antes de submeter os dados às redes neurais para realizar os testes, é necessário normalizá-los, isto é, deixar todos dentro de um mesmo intervalo. A normalização faz com que todas as entradas tenham a mesma relevância para o treinamento, ou seja, as variáveis, mesmo que tenham ordens de grandeza diferentes, terão a mesma relevância nos experimentos. Outro efeito da normalização é deixar os dados dentro dos limites da função de ativação (VALENÇA, 2005).

O intervalo utilizado na normalização depende da função de ativação da camada de saída. Por exemplo, caso função de ativação seja a função sigmoide logística, os valores tem que ser limitados entre 0 e 1, logo seus valores devem ser normalizados entre 0,10 e 0,90 ou entre 0,15 e 0,85. Caso a função de ativação seja a tangente hiperbólica os valores são normalizados entre -0,90 e 0,90 ou entre -0,85 e 0,85. Na normalização os dados não devem ser re-escalados nos extremos da função, pois algumas funções quando trabalham com valores que estão em seus extremos se aproximam de zero.

De acordo com Valença (2005), os tipos de normalização mais frequentemente utilizados são:

- transformação linear:

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

- normalização estatística:

$$y = \frac{x_i - \bar{x}}{s} \quad (3.2)$$

- normalização simples:

$$y = \frac{x_i}{x_{max}} \quad (3.3)$$

Temos: y é o valor normalizado; x_i é o valor original; x_{max} é o valor máximo da variável; x_{min} é o valor mínimo da variável; \bar{x} é o valor médio da variável e s o desvio padrão.

Para a base de dados do Parque I utilizada nos experimentos deste trabalho foram utilizados os limites $[-0,85$ e $0,85]$, logo assume-se o uso da função tangente hiperbólica, para normalizar os dados utilizando a transformação linear.

3.2.2 Defasagem dos Dados

A previsão é realizada utilizando valores medidos em instantes de tempo anteriores ao atual. Assim, após a normalização deve-se realizar a defasagem dos dados de potência eólica, para que assim se tenha como entradas da rede os dados em um instante de tempo anterior ao do dado atual.

A defasagem constitui um processo que torna um dado medido em uma época t , ocupando uma posição i do vetor de dados, presente na época $t - 1$, ocupando a posição $i + 1$ do vetor. A tabela 1 mostra como foi realizada a defasagem dos valores presentes na base de dados do Parque I.

Tabela 1 – Representação de defasagem dos valores presentes na base de dados do Parque I

| t | 5 | 4 | 3 | 2 | 1 | 0 |
|-----|---|---|---|---|---|---|
| t-1 | 4 | 3 | 2 | 1 | 0 | |
| t-2 | 3 | 2 | 1 | 0 | | |
| t-3 | 2 | 1 | 0 | | | |
| t-4 | 1 | 0 | | | | |

3.3 Arquiteturas de Redes Neurais na Previsão de Potência Eólica

3.3.1 *Multi-Layer Perceptron*

Para a configuração das diversas arquiteturas de RNAs utilizadas nos experimentos propostos neste trabalho foi utilizado o software *Weka* (WEKA, 2008).

O primeiro tipo de RNA utilizado nos experimentos foi a MLP. Para a configuração dos parâmetros da MLP utilizou-se o *Weka 3.7.8* através do módulo *Explorer* (BOUCKAERT et al., 2008), na aba *Classify*, selecionando a rede neural MLP na opção *Functions*.

Após selecionar a MLP, foi necessário configurar alguns parâmetros para a correta execução dos testes:

- *learningRate* - taxa de aprendizagem, foi configurada com o valor 0,3;
- *seed* - semente, inicial configurada com o valor 1, mas é preciso alterar o seu valor antes de cada simulação, caso contrário, todas as simulações vão apresentar os mesmo resultados;
- *validationSetSize* - define a porcentagem dos dados submetidos ao treinamento, e a porcentagem que será utilizada na validação cruzada.

Como foi definido nos objetivos, durante os experimentos ocorreram variações no horizonte de entrada. Primeiro foram feitas 30 simulações com 6 entradas. Como os valores presentes na base de dados foram medidos em intervalos de 30 minutos, este horizonte totaliza 3 horas, ou seja, com essa configuração foram usados os dados de 3 horas para prever as próximas 3 horas.

Após as simulações com 6 entradas foram feitas 30 simulações com 12 entradas, ou seja, utilizou-se valores de potências medidas durante 6 horas para prever as próximas 3 horas.

Em seguida, foi utilizado um horizonte de 24 entradas, ou seja, 12 horas. Após as simulações com 24 entradas foram realizadas simulações com 48 entradas. É importante lembrar que o valor da semente era alterado antes de cada simulação, e voltava a ser 1 quando era alterado o horizonte de entrada.

O *software* utilizado nos experimentos gerou os seguintes resultados:

- *Correlation coefficient*
- *Mean absolute error* - EMA
- *Root mean squared error* - REMQ

- *Relative absolute error* - ERA
- *Root relative squared error* - RERQ

Para os objetivos proposto neste trabalho o valor de interesse entre os valores gerados nas simulações é a REMQ, logo para cada simulação será salvo este valor. Na seção 3.3.4 é mostrado como calcular o valor da REMQ.

3.3.2 *Radial Basis Function*

Após executar os testes com a rede MLP, o passo seguinte foi realizar os experimentos com a rede RBF. Antes foi necessário configurá-la, com o auxílio da mesma ferramenta utilizada para os testes com MLP, *Weka 3.7.8*.

Para configurar a rede RBF na ferramenta utilizada é preciso seguir os mesmos passos que foram seguidos para a outra rede, e na opção *Functions* selecionar *RBFNetwork*. Os parâmetros configuráveis da rede RBF na ferramenta utilizada são:

- *clusteringSeed* - Valor da semente: inicialmente seu valor é foi definido em 1. Assim como na MLP esse valor deve ser alterado antes de cada iteração;
- *numCluster* - define o número de *clusters* nos quais os dados da entrada serão divididos. Seu valor foi definido em 30.

Antes da definição dos valores destes parâmetros foram realizados testes alterando seus valores. Nesses testes verificou-se que quanto maior o número de *clusters*, menos as taxas de erro das simulações se repetiam. Logo, como seriam realizadas 30 simulações o número de *clusters* foi definido em 30.

Para os testes com a rede RBF, a variação de entradas foi a mesma que ocorreu com a MLP. Primeiro foi utilizada uma arquitetura com 6 entradas e foram executadas 30 simulações, sempre alterando, antes da execução de cada uma, o valor da semente. Para cada execução foi salvo o valor da REMQ. Após executar as 30 simulações a arquitetura da rede foi alterada para 12 variáveis de entrada. O mesmo processo foi executado, ao final das simulações a arquitetura foi alterada para 24 entradas. Os mesmos testes foram executados, e após 30 simulações utilizou-se 48 entradas e os mesmos passos foram executados.

A ferramenta apresenta após cada simulação os mesmos resultados que apresentou para MLP. Para os objetivos deste trabalho os valores a serem salvos foram as REMQs.

3.3.3 *Support Vector Machine*

Após os testes com RBF foram realizados testes com SVM. Para sua configuração foi utilizada a mesma ferramenta que foi utilizada nas outras redes.

Para configurar a rede SVM acessamos a opção *Functions* e selecionamos *LibSVM*. Antes realizar os testes foi necessário configurar alguns parâmetros:

- *SVMType* - define o tipo de SVM a ser usado: para este trabalho foi escolhido o tipo *nu-SVR*
- *kernelType* - define o tipo de função que será usado: para este trabalho foi selecionada a função *radial basis function*;
- *seed* - define o valor da semente: inicialmente foi configurada com o valor 1.

Para definir os valores dos parâmetros, foram realizados testes com diversos valores aplicados nesses parâmetros.

O *SVMType* foi configurado com *nu-SVR* pois apresentou resultados mais satisfatórios do que o *epsilon-SVR*. O *kernelType* foi configurado com a função RBF por causa do baixo número de parâmetros (dois) que influenciam o resultado do aprendizado. No *kernel* polinomial são usados mais parâmetros que o RBF: C (custo), γ (*gamma*) e o grau (*degree*). Se C assumisse 11 valores, γ assumisse 10 no *kernel* RBF, seria uma combinação de 110 valores, adicionando-se o grau e atribuindo-lhe 5 valores, a combinação iria para 550, aumentando consideravelmente o desempenho e complexidade do classificador. Os valores destes parâmetro são definidos pela ferramenta (JUNIOR, 2010).

O procedimento para os experimentos é igual ao realizado com as outras redes. Primeiro realizou-se os testes com 6 variáveis de entrada. Foram feitas 30 simulações salvando para cada uma o valor da REMQ. Terminados os testes com 6 variáveis, foram realizados testes com 12 variáveis, salvando para estas simulações o valor da REMQ. Os mesmos passos foram executados com 24 e 48 variáveis de entrada.

Após realizar todas as simulações com RNAs, foram realizados testes estatísticos com os valores de REMQ gerados. Este teste é descrito na seção 3.4. Porém, antes é necessário mostrar o procedimento utilizado para os testes.

3.3.4 Validação Cruzada

Segundo Valença (2009) durante o treinamento um fator importante é o critério utilizado para encerrá-lo. Alguns estudos utilizam critérios como número máximo de ciclos ou erro mínimo. Mas em algumas aplicações esses critérios não são satisfatórios pois não levam em conta o processo iterativo.

Escolher o critério de parada adequado é importante para evitar dois problemas que podem ocorrer durante o treinamento: *underfitting* e *overfitting*.

O *underfitting* acontece quando o treinamento é parado precocemente, assim a rede com baixa capacidade de generalização. Seria como se um botânico classificasse qualquer objeto como árvore apenas por ser verde.

Já o *overfitting* ocorre quando a rede é treinada em excesso, o que a deixa superajustada. Com isso a rede passa a memorizar perdendo sua capacidade de generalização. Voltando ao exemplo do botânico, seria como se ele classificasse uma árvore como não sendo uma árvore apenas por ter menos folhas que os exemplos vistos anteriormente.

A validação cruzada utiliza um conjunto de dados (normalmente 25% da amostra) para informar quando devemos para o treinamento. Esses dados são utilizados para calcular a função erro utilizada para interromper o treinamento quando esta atingir seu valor mínimo ou de um ciclo para o outro seu valor ficar em estacionário dentro de uma dada precisão.

A validação cruzada é realizada da seguinte forma:

- Apresenta-se a rede neural o conjunto de treinamento utilizado para ajustar os pesos. Quando se conclui a apresentação de todos os exemplos diz-se que se completou um ciclo.
- A cada ciclo após o reajuste dos pesos, o treinamento é parado e o conjunto de validação cruzada é apresentado para que se possa calcular o erro.
- O processo é repetido a cada ciclo de modo que se possa traçar um gráfico onde se coloca na horizontal o número de ciclos e na vertical o valor do erro.
- O treinamento é interrompido quando o valor do erro, para o conjunto de validação cruzada, após atingir seu valor mínimo, começa a crescer ou quando o valor do erro de um ciclo para o outro ficar estacionário dentro de uma dada precisão.

A função mais comum de ser utilizada é a função de error médio quadrático:

$$EMQ = \frac{1}{N} \sum_{p=0=1}^N (Z_p - Z_o)^2 \quad (3.4)$$

Mas nos experimentos realizados neste trabalho foi utilizado o valor da raiz do erro médio quadrático (REMQ), que mede a amplitude do erro:

$$REMQ = \sqrt{\frac{1}{N} \sum_{p=0=1}^N (Z_p - Z_o)^2} \quad (3.5)$$

Nas equações 3.4 e 3.5 temos os seguintes parâmetros:

- N - número de valores previstos;

- Z_p - valor previsto;
- Z_o - valor ocorrido.

3.4 Testes Estatísticos

Após 30 simulações com cada configuração de rede neural, foram realizados testes estatísticos para verificar qual método apresenta o melhor desempenho na previsão de potência eólica ou os resultados podem ser considerados estatisticamente iguais. Para realização dos teste utilizou-se o *software* matemático R, que contém as implementações dos testes utilizados neste trabalho. Este *software* utiliza como padrão um nível de significância (α) previamente definido com o valor 0,05.

Entre os teste disponíveis na literatura se encontram os testes *T-Student* e *Wilcoxon*. Antes de aplicar o *T-Student* é necessário aplicar o teste de *Shapiro-Wilk* que analisa se os dados são normalmente distribuídos. Após aplicar o teste *Shapiro-Wilk*, é aplicado o Teste F, para verificar se a variância é originada da mesma população. Caso os dados passem nos dois testes é aplicado então o *T-Student*. Caso eles não passem é aplicado o teste não paramétrico das Somas dos Postos de *Wilcoxon*.

3.4.1 Teste de *Shapiro-Wilk*

Este teste tem o objetivo de verificar se as amostras são normalmente distribuídas e se possuem variância estatisticamente iguais. Se essas premissas forem verdadeiras, pode-se aplicar um teste paramétrico, caso contrário, um teste não paramétrico deve ser aplicado.

Assim, antes de executar esse teste é necessário formular duas hipóteses:

- Hipótese nula: as amostras são normalmente distribuídas.
- Hipótese alternativa: as amostras não são normalmente distribuídas.

Após executar o teste para cada modelo, deve-se analisar o *p-value* e compará-lo com o nível de significância. Caso o *p-value* seja menor que α a hipótese nula é rejeitada e a amostra não é considerada normalmente distribuída.

3.4.2 Teste F

Após o teste de *Shapiro-Wilk*, deve-se aplicar o Teste F, que tem o objetivo de verificar os dados de dois conjuntos são originados da mesma população.

Antes de aplicar o teste é preciso formular duas hipóteses:

- Hipótese nula: as amostras possuem variâncias estatisticamente iguais.
- Hipótese alternativa: as amostras não possuem variâncias estatisticamente iguais.

Após executar o teste, o valor do *p-value* é analisado. Se for menor que o nível de significância, a hipótese nula é rejeitada.

3.4.3 Teste *t-Student*

O *T-Student* é um teste paramétrico. Assim, para que seja aplicado é necessário que os modelos a serem testados sejam normalmente distribuídos e possuam variâncias estatisticamente iguais.

Assim como nos outros testes, antes de aplicar o *T-Student* é necessário formular as hipóteses nula e alternativa:

- Hipótese nula: as médias das amostras são estatisticamente iguais.
- Hipótese alternativa: as médias das amostras não são estatisticamente iguais.

Para determinar qual das hipóteses é a verdadeira compara-se o *p-value* com o nível de significância. Caso o *p-value* seja menor a hipótese nula é rejeitada e o modelo com maior média é considerado melhor.

3.4.4 Teste das Somas dos Postos de *Wilcoxon*

O teste de *Wilcoxon* é um teste não paramétrico utilizado para análise de dados os valores não são normalmente distribuídos ou não possuem variâncias estatisticamente iguais. Esse teste faz sua análise baseado na mediana de cada uma das amostras analisadas.

Este teste apresenta duas hipóteses:

- Hipótese nula: os modelos comparados tem distribuição simétrica.
- Hipótese alternativa: há uma diferença significativa entre os modelos analisados.

Para definir qual das hipóteses é a verdadeira usa-se como referência a variável *p-value*. Antes de executar os testes deve-se definir um nível de significância, neste trabalho foi adotado o valor 0,05. Após executar o teste compara-se o *p-value* com nível de significância, se o *p-value* for maior quer dizer que a hipótese nula é a verdadeira, caso contrário é a hipótese alternativa.

4 Resultados

Este capítulo apresenta os resultados obtidos após a aplicação de todos os passos descritos no capítulo anterior.

Para a configuração das arquiteturas utilizadas neste trabalho foi utilizado o *software Weka 3.7.8*, através do módulo *Explorer*. Foram utilizados como valores iniciais dos parâmetros, das 3 três arquiteturas, aqueles definidos nas seções 3.3.1, 3.3.2 e 3.3.3.

Foram executadas 30 simulações para cada configuração (tipo de RNA e quantidade de entradas) para realizar previsões a cada intervalo de 30 minutos totalizando três horas à frente. De cada simulação foi salvo o valor da raiz do erro médio quadrático (REMQ), a fim de submetê-los aos testes estatístico citados anteriormente.

As próximas seções apresentam os resultados dessas simulações e as aplicações dos testes estatísticos.

4.1 Comparativos e Testes estatísticos entre MLP, RBF e SVM

4.1.1 Comparação da Raiz do Erro Médio Quadrático

Esta seção apresenta os valores das REMQs geradas a partir das simulações, isto não define a rede mais adequada. A rede mais adequada só será definida com a aplicação dos testes estatísticos, apresentado na seção 3.4.

Os valores de REMQ gerados a partir das simulações com 6 valores podem ser vistos na Figura 20. Nela pode-se observar que a rede MLP apresentou valores maiores de REMQ, assim a variação deste valor foi maior para esta rede. Já as redes RBF e SVM, apresentaram valores muito semelhantes. A rede RBF apresentou pouca variação, enquanto o modelo SVM não apresentou variação alguma no valor da REMQ.

As simulações feitas com 12 valores de entradas tem seus valores de REMQ apresentados na Figura 21. A partir deste gráfico percebe-se um comportamento semelhante ao das simulações com 6 valores. A MLP apresentou maior variação e valores maiores da REMQ. No caso das redes RBF e SVM apresentaram valores próximos, com a RBF variando muito pouco e a SVM apresentando valor constante, porém, menor que as outras redes.

Analisando a Figura 22 podemos observar uma pequena mudança nos valores de REMQ das simulações com 24 valores em relação as simulações comentadas anteriormente. A MLP apresentou as mesmas características das configurações anteriores: maior variação no valor da REMQ e maior dessa taxa de erro. A mudança em relação às Figuras 20 e 21

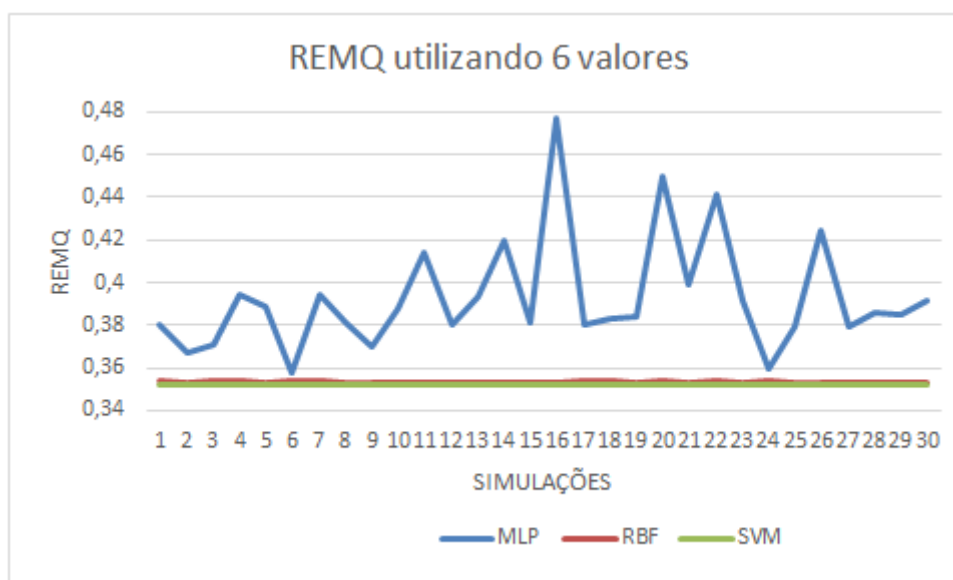


Figura 20 – Comparação da Raiz do EMQ com 6 valores

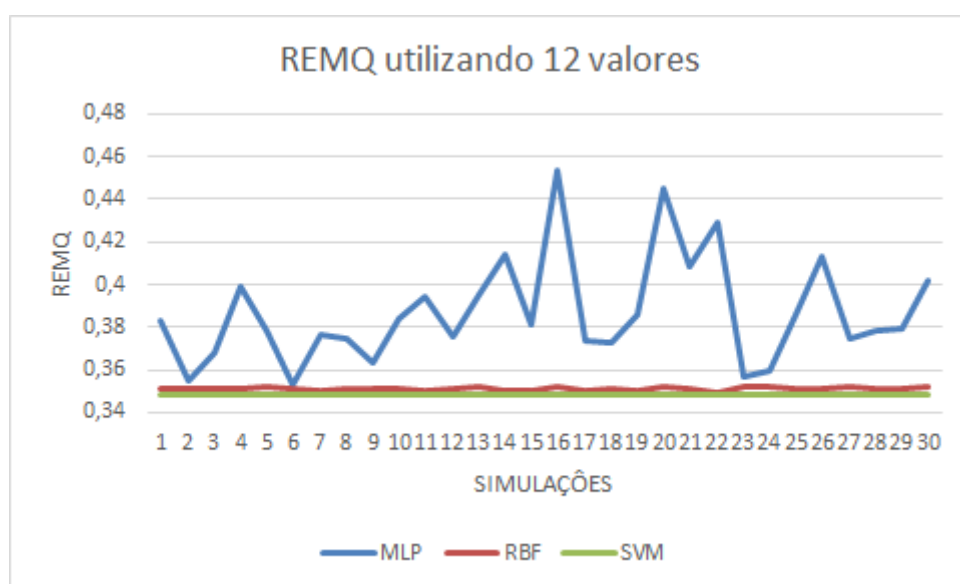


Figura 21 – Comparação da Raiz do EMQ com 12 valores

está no comportamento das redes RBF e SVM. Elas apresentaram valores distintos, com a primeira apresentando pouco variação e a segunda valor constante e menor em relação aos valores de REMQ apresentados pelas outras redes.

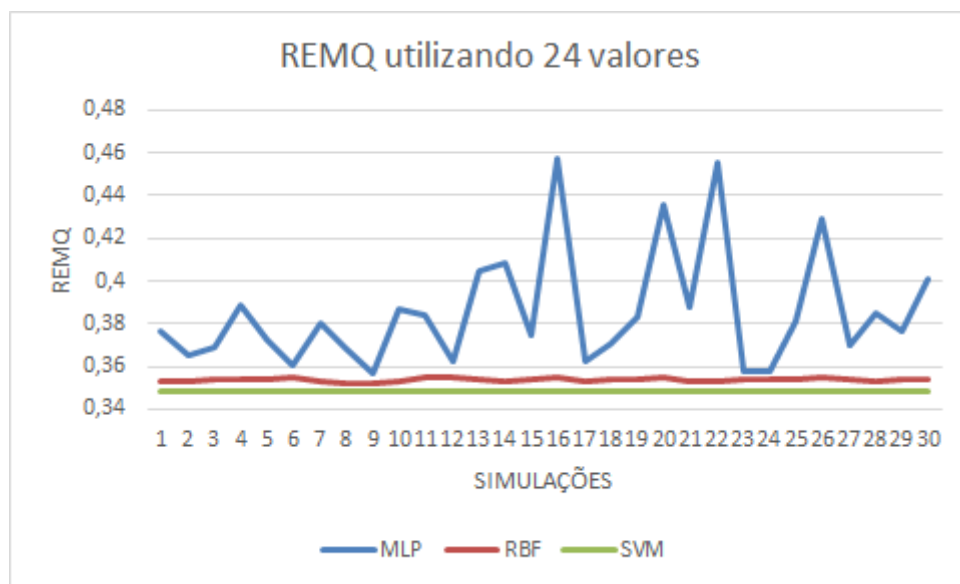


Figura 22 – Comparação da Raiz do EMQ com 24 valores

Por fim, as simulações com 48 entradas tem seus valores de REMQ apresentados na Figura 23. Nesta configuração a rede RBF apresentou os valores maiores de REMQ. A MLP apresentou uma variação em relação à sua utilização em outras configurações. A rede SVM apresentou novamente valor constante, porém menor que as outras redes.

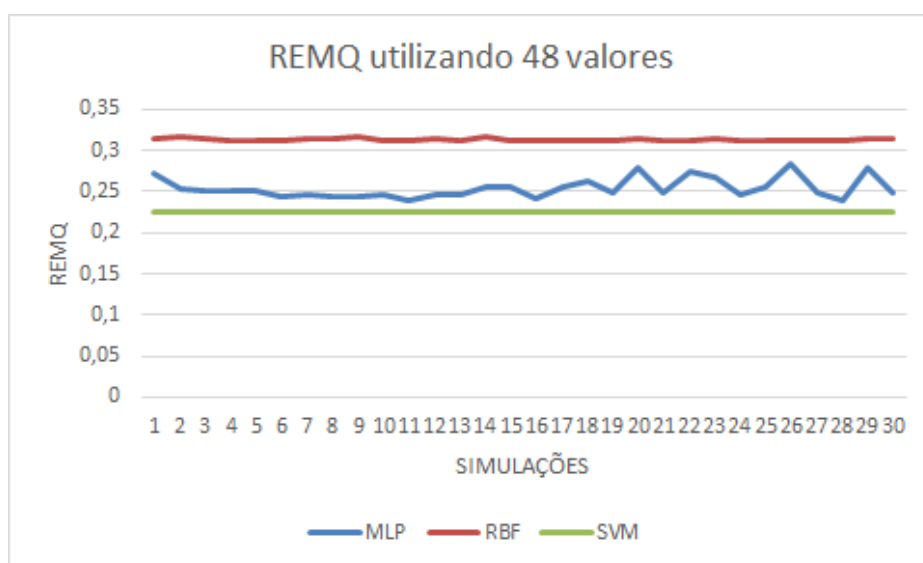


Figura 23 – Comparação da Raiz do EMQ com 48 valores

4.1.2 Testes Estatísticos

Após aplicação do teste de *Shapiro-Wilk* e Teste F, verificou-se que não poderia ser aplicado o teste *T-Student*, pois as amostras - resultantes dos testes com os 4 horizontes de entrada - não passaram nos dois primeiros. Desse modo para a análise dos resultado foi utilizado o teste não paramétrico de *Wilcoxon*.

4.1.2.1 Arquiteturas com 6 valores de entrada

O resultado das simulações com 6 valores de entradas é apresentado na Tabela 2. Nela podemos ver a raiz do EMQ para as 30 simulações de cada topologia.

Tabela 2 – Representação da média das taxas de erro após 30 simulações com cada arquitetura utilizando 6 variáveis.

| Topologia da rede neural | Raiz do EMQ (30 simulações) |
|--------------------------|-----------------------------|
| MLP | 0,393156667 |
| RBF | 0,353256111 |
| SVM | 0,351583333 |

A partir dos resultados das simulações foi aplicado o teste da Soma dos Postos de *Wilcoxon*. Para cada conjunto REMQs, obtidos a partir das simulações com as arquitetura MLP, RBF e SVM, foi realizado o teste não-paramétrico de *Wilcoxon*, para verificar se os modelos são estatisticamente distintos. Este teste foi aplicado utilizando o *software R*, com o nível de significância estabelecido em 0,05. A Tabela 3 mostra os resultados destes testes.

Tabela 3 – Resultados do teste da Soma dos Postos de *Wilcoxon* para 6 valores de entrada.

| RNAs testadas | <i>p-value</i> |
|---------------|-------------------------|
| MLPxRBF | 0,464 |
| MLPxSVM | $4,574 \times 10^{-12}$ |
| SVMxRBF | $4,552 \times 10^{-12}$ |

A partir da Tabela 3 pode-se concluir da comparação entre MLP e RBF, que ambas são estatisticamente iguais pois o *p-value* foi maior que o nível de significância (0,05), logo, pelo teste de Wilcoxon pode-se utilizar qualquer uma das duas.

Por outro lado, a comparação entre MLP e SVM resultou em um *p-value* menor que o nível de significância. O que nos leva a escolher o modelo que apresentou menor taxa de erro, logo, a opção mais adequada é a rede SVM que apresentou REMQ no valor de 0,351583333.

Para a comparação entre SVM e RBF ocorre o mesmo, o *p-value* calculado foi menor que 0,05 e a rede SVM apresenta a menor taxa de erro, sendo escolhida a mais adequada para previsão com 6 valores.

4.1.2.2 Arquiteturas com 12 valores de entrada

Na Tabela 4 podemos ver as REMQs resultantes das 30 simulações com 12 valores de entradas para cada topologia.

Tabela 4 – Representação da média das taxas de erro após 30 simulações com cada arquitetura utilizando 12 variáveis.

| Topologia da rede neural | Raiz do EMQ (30 simulações) |
|--------------------------|-----------------------------|
| MLP | 0,387228889 |
| RBF | 0,351083333 |
| SVM | 0,348683333 |

Assim como foi feito com os resultados das simulações com 6 valores, os resultados das simulações com 12 valores foram utilizados para a aplicação de teste de *Wilcoxon* e assim determinar se os modelos são estatisticamente distintos. Os resultados do teste de *Wilcoxon* são mostrados na Tabela 5.

Tabela 5 – Resultados do teste da Soma dos Postos de *Wilcoxon* para 12 valores de entrada.

| RNAs testadas | <i>p-value</i> |
|---------------|-------------------------|
| MLPxRBF | 0,4731 |
| MLPxSVM | $4,574 \times 10^{-12}$ |
| SVMxRBF | $4,563 \times 10^{-12}$ |

A partir da Tabela 5 concluímos que as redes MLP e RBF são estatisticamente iguais. Já na comparação de ambas com a SVM concluímos que esta é mais adequada do que as outras pois no teste da SVM com a MLP e RBF o *p-value* foi menor que 0,05, permitindo escolher a SVM por apresentar menor taxa de erro.

4.1.2.3 Arquiteturas com 24 valores de entrada

O mesmo procedimento realizado nas arquiteturas com 6 e 12 valores de entrada foi realizado para a arquitetura com 24 entradas. Os resultados das 30 simulações, mostrados na Tabela 6, foram utilizados para a aplicação do teste de *Wilcoxon*, onde os resultados deste último teste são apresentados na Tabela 7.

Com o auxílio das tabelas 6 e 7 concluímos que a rede SVM é a mais adequada para previsão de potência eólica com 24 entradas. Pois, o teste realizado com MLP e RBF

Tabela 6 – Representação da média das taxas de erro após 30 simulações com cada arquitetura utilizando 24 variáveis.

| Topologia da rede neural | Raiz do EMQ (30 simulações) |
|--------------------------|-----------------------------|
| MLP | 0,385669444 |
| RBF | 0,353782222 |
| SVM | 0,348083333 |

Tabela 7 – Resultados do teste da Soma dos Postos de *Wilcoxon* para 24 valores de entrada.

| RNAs testadas | <i>p-value</i> |
|---------------|-------------------------|
| MLPxRBF | 0,569 |
| MLPxSVM | $4,574 \times 10^{-12}$ |
| SVMxRBF | $4,563 \times 10^{-12}$ |

resultou em um *p-value* maior que o nível de significância definido em 0,05, sendo essas duas RNAs estatisticamente iguais. Já nos testes realizados comparando SVM com MLP e RBF, o *p-value* encontrado foi menor que 0,05, sendo a opção mais adequada aquela que apresenta a menor taxa de erro, logo, SVM.

4.1.2.4 Arquiteturas com 48 valores de entrada

Por fim, na arquitetura com 48 valores foram realizados os mesmos procedimentos que nas outras arquiteturas, executando-se 30 simulações com cada topologia, guardando sempre o valor da REMQ, os resultados são mostrados na Tabela 8. Em seguida, foi aplicado o teste de *Wilcoxon* a fim de encontrar a topologia mais adequada a previsão com 48 valores, o resultado desse teste é mostrado na Tabela 9.

Tabela 8 – Representação da média das taxas de erro após 30 simulações com cada arquitetura utilizando 48 variáveis.

| Topologia da rede neural | Raiz do EMQ (30 simulações) |
|--------------------------|-----------------------------|
| MLP | 0,254461111 |
| RBF | 0,31354 |
| SVM | 0,22535 |

Tabela 9 – Resultados do teste da Soma dos Postos de *Wilcoxon* para 48 valores de entrada.

| RNAs testadas | <i>p-value</i> |
|---------------|-------------------------|
| MLPxRBF | 0,2704 |
| MLPxSVM | $4,574 \times 10^{-12}$ |
| SVMxRBF | $4,534 \times 10^{-12}$ |

Assim como nas outras arquiteturas a rede SVM foi determinada como a mais adequada para a previsão utilizando 48 entradas. Pois, o teste de *Wilcoxon* para MLP e RBF resultou em um *p-value* maior que o nível de significância, e o teste entre SVM as outras duas redes resultou em um *p-value* menor que 0,05. Sendo assim, a rede escolhida como a mais adequada é aquela que apresentar a menor taxa de erro. Como podemos ver na Tabela 8 essa rede é a SVM.

Após encontrar a rede mais adequada para cada horizonte de entrada, é preciso determinar a mais adequada entre elas para a previsão de potência eólica.

A Tabela 10 mostra as taxas de erro para cada rede determinada como a mais adequada para cada horizonte.

Tabela 10 – Representação das taxas de erro de cada rede escolhida como mais adequada para horizonte de entrada.

| Horizonte | RNA | REMQ |
|-------------|-----|-------------|
| 6 entradas | SVM | 0,351583 |
| 12 entradas | SVM | 0,348683333 |
| 24 entradas | SVM | 0,348083333 |
| 48 entradas | SVM | 0,22535 |

Observando a Tabela 10 podemos notar que a rede SVM com 48 entradas apresentou menor taxa de erro. Para determinar qual das redes presentes nesta tabela é a mais adequada, seria necessário aplicar testes estatísticos. No entanto, como pode ser visto na seção 4.1.1, a rede SVM não apresenta variação na taxa, para as 30 simulações de cada horizonte. Não sendo possível aplicar tais testes.

Sendo assim, conclui-se que as quatro configurações de SVM são estatisticamente iguais, pois apresentaram taxas de erro semelhantes durante os experimentos realizados. Assim, qualquer uma delas pode ser utilizada para previsão de potência eólica, pois apresentarão resultados semelhantes quando realizarem as previsões.

5 Considerações Finais

5.1 Conclusões

Este trabalho teve como objetivo selecionar a arquitetura de rede neural mais adequada à previsão de potência eólica três horas à frente. As redes neurais utilizadas para realizar a previsão foram MLP, RBF e SVM. Nos experimentos realizados as configurações das redes variaram em relação à quantidade de valores de entrada. Assim foram utilizadas para cada topologia arquiteturas com 6, 12, 24 e 48 valores de entrada.

Foram utilizadas os módulos de redes neurais disponíveis no *Weka* 3.7.8. De posse das ferramentas e do dados foram realizados várias simulações com as configurações de cada definidas anteriormente. Para cada simulação foi salvo o valor da raiz do erro médio quadrático.

Os testes estatísticos realizados permitiram analisar o desempenho das topologias utilizadas. Essa análise determinou que como os resultados foram similares as redes SVM para as arquiteturas testadas são estatisticamente iguais. Assim, as quatro configurações apresentam valores semelhantes para previsão de potência eólica.

5.2 Trabalhos Futuros

Após concluir os experimentos e analisar seus resultados, espera-se em outros estudos realizar os experimentos propostos neste trabalho com cenários diferente daqueles que foram objetos de estudo. Sendo assim, para trabalhos futuros pretende-se estudar a previsão de potência eólica sob novas perspectivas, tais como:

- Novas topologias de redes neurais para serem comparadas àquelas que foram objetos de estudo no trabalho atual.
- Novas ferramentas, que sejam mais flexíveis no que diz respeito a configuração das redes neurais, visto que a ferramenta utilizada neste trabalho não permite por exemplo atribuir qualquer quantidade de neurônios na camada escondida.
- Novas arquiteturas: neste trabalho utilizou-se 4 horizontes de entrada (6, 12, 24 e 48), pretende-se no futuro testar outros horizontes de entrada, assim como variar a quantidade de neurônios presentes na camada escondida.

-
- Diferentes intervalos de previsão: neste trabalho foi utilizada previsão a cada 30 minutos. Pretende-se testar outros intervalos.
 - Variar o tempo a ser previsto: pretende-se estudar a previsão com alcance diferentes do utilizado no trabalho atual que foi de 3 horas à frente.

Referências

- ALIAGA, D. K. Redes neurais artificiais rbf. PGEAS - UFSC, 2010. Acessado em 02/10/2015.
- AMARANTE, O. A. et al. Atlas do potencial eólico brasileiro. In: *Atlas do potencial eólico brasileiro*. [S.l.]: Ministerio de Minas e EnergiaEletrobras, 2001.
- ANDREOLA, R. Support vector machines na classificação de imagens hiperespectrais. 2009. Acessado em 23/08/2015.
- ANEEL. *Atlas da Energia Elétrica do Brasil*. TDA Comunicação, 2005. Acesso em 18/08/2015. Disponível em: <[http://www.aneel.gov.br/aplicacoes/atlas/pdf/06-energia_eolica\(3\).pdf](http://www.aneel.gov.br/aplicacoes/atlas/pdf/06-energia_eolica(3).pdf)>.
- ANEEL. *Atlas da Energia Elétrica do Brasil*. 3. ed. TDA Comunicação, 2008. Acesso em 18/08/2015. Disponível em: <<http://www.aneel.gov.br/arquivos/PDF/atlas3ed.pdf>>.
- BOLDUAN, P. *Old windmill (1880 town, South Dakota, USA)*. 2000. Acessado em 27/10/2015. Disponível em: <https://commons.wikimedia.org/wiki/File:Old_Windmill.jpg>.
- BOUCKAERT, R. et al. *Weka manual*. 2008. Acessado em 15/08/2015.
- CARDOSO, M. R. A. Análise de séries temporais em epidemiologia: uma introdução sobre os aspectos metodológicos. *Rev. bras. epidemiol*, SciELO Public Health, v. 4, n. 3, 2001. Acessado em 13/10/2015.
- COELHO, L. d. S.; SANTOS, A. A. P.; JR, N. C. A. d. C. Podemos prever a taxa de cambio brasileira? evidência empírica utilizando inteligência computacional e modelos econométricos. *Gestão & Produção*, v. 15, n. 3, 2008. Acessado em 21/11/2015.
- EHLERS, R. S. Análise de séries temporais. *Laboratório de Estatística e Geoinformação. Universidade Federal do Paraná*, 2007. Acessado em 23/10/2015.
- FERREIRA, A. A. d. A. *Sistema de produção de energia eólica*. Tese (Doutorado) — Universidade do Porto, 2011. Acessado em 27/10/2015.
- FONSECA, I. d. S. A. d. et al. Manutenção de sistemas de geração de energia renovável eólica através de redes ip. 2012.
- HECKL, W. M. *Deutsches Museum*. 2015. Acessado em 27/10/2015. Disponível em: <<http://www.deutsches-museum.de/>>.
- JUNIOR, P. M. d. O. *Máquina de Vetores Suporte: estudo e análise de parâmetros para otimização de resultado*. Trabalho de conclusÃo de curso — Universidade Federal de Pernambuco, 2010. Acessado em 27/11/2015.

- LEITE, V. *Uma análise da classificação de litologias utilizando SVM, MLP e métodos Ensemble*. Tese (Doutorado) — Dissertação de Mestrado. Departamento de Informática. Rio de Janeiro: Pontifícia Universidade Católica do Rio de Janeiro, 2012. Acessado em 10/11/2015.
- LUDERMIR, T. B. et al. *Redes neurais artificiais: teoria e aplicações*. LTC Editora, 2007. Acessado em 18/08/2015.
- MME, M. d. M. e. E. *Programa de Incentivo às Fontes Alternativas de Energia Elétrica*. 2004. Acessado em 02/11/2015. Disponível em: <<http://www.mme.gov.br/programas/proinfa/>>.
- MOLENAAR, D.-P. *Cost-effective design and operation of variable speed wind turbines*. TU Delft, Delft University of Technology, 2003. Acessado em 03/11/2015.
- PAZ, W. V. d. *PREVISÃO DE GERAÇÃO DE POTÊNCIA EÓLICA UTILIZANDO REDES NEURAIIS ARTIFICIAIS*. Trabalho de conclusão de curso — Universidade de Pernambuco, 2014. Acessado em 19/11/2015.
- RODRIGUES, G. C. *Utilização de Redes Neurais para Previsão de Ventos no Horizonte de 24 Horas*. Tese (Doutorado) — UNIVERSIDADE FEDERAL DO RIO DE JANEIRO, 2007. Acessado em 17/09/2015.
- RODRIGUES, R. C. B. et al. *Máquinas de vetores de suporte aplicadas à classificação de defeitos em couro bovino*. Trabalho de conclusão de curso — Universidade Católica Dom Bosco, 2007. Acessado em 20/11/2015.
- SALLES, M. B. d. C. *Modelagem e análises de geradores eólicos de velocidade variável conectados em sistemas de energia elétrica*. Tese (Doutorado) — Universidade de São Paulo, 2009. Acessado em 20/10/2015.
- SANTOS, E. M. dos. *Teoria e Aplicação de Support Vector Machines à Aprendizagem e Reconhecimento de Objetos Baseado na Aparência*. Tese (Doutorado) — Universidade Federal da Paraíba, 2002. Acessado em 15/11/2015.
- SILVA, I. N. da et al. *Redes neurais artificiais para engenharia e ciências aplicadas curso prático*. Artliber, SciELO Brasil, 2010.
- TABLADA, L. G. N. *Utilizando Redes Neurais para a Previsão de Velocidade do Vento*. Trabalho de conclusão de curso — Universidade de Pernambuco, 2010. Acessado em 27/08/2015.
- VALENÇA, M. *Aplicando redes neurais: Um guia completo*. Livro Rápido, Olinda, 2005. Acessado em 18/10/2015.
- VALENÇA, M. *Fundamentos das redes neurais: exemplos em java*. Livro Rápido, Olinda, 2009. Acessado em 18/08/2015.
- VITERBO, J. C. *Geração de energia elétrica a partir da fonte eólica offshore*. Tese (Doutorado) — Universidade de São Paulo, 2008. Acessado em 20/10/2015.
- WEKA. *Data Mining Software in Java*. 2008. Acessado em 15/08/2015. Disponível em: <<http://www.cs.waikato.ac.nz/ml/weka/>>.