



**UNIVERSIDADE FEDERAL DO CEARÁ**  
**CENTRO DE TECNOLOGIA**  
**DEPARTAMENTO DE ENGENHARIA MECÂNICA**  
**CURSO DE ENGENHARIA MECÂNICA**

**JEFFERSON LEMOS FERNANDES**

**PREVISÃO DE IRRADIAÇÃO SOLAR ATRAVÉS DE UM MODELO DE REDE  
NEURAL OTIMIZADA COM O ALGORITMO DE BROYDEN-FLETCHER-  
GOLDFARB-SHANNO**

**FORTALEZA**

**2016**

**JEFFERSON LEMOS FERNANDES**

**PREVISÃO DE IRRADIAÇÃO SOLAR ATRAVÉS DE UM MODELO DE REDE  
NEURAL OTIMIZADA COM O ALGORITMO DE BROYDEN-FLETCHER-  
GOLDFARB-SHANNO**

Trabalho de conclusão de curso apresentado ao curso de Engenharia Mecânica do Departamento de Engenharia Mecânica da Universidade Federal do Ceará, como requisito parcial para a obtenção do título de Engenheiro Mecânico.

Orientador: Prof. Dr. Paulo Alexandre Costa Rocha.

**FORTALEZA**

**2016**

Dados Internacionais de Catalogação na Publicação  
Universidade Federal do Ceará  
Biblioteca de Ciências e Tecnologia

---

F399p

Fernandes, Jefferson Lemos.

Previsão de irradiação solar através de um modelo de rede neural otimizada com o algoritmo de Broyden-Fletcher-Goldfarb-Shanno / Jefferson Lemos Fernandes. – 2016.  
52 f.: il. ; color.

Monografia (Graduação) – Universidade Federal do Ceará, Centro de Tecnologia,  
Departamento de Engenharia Mecânica, Curso de Engenharia Mecânica, Fortaleza, 2016.  
Orientação: Prof. Dr. Paulo Alexandre Costa Rocha.

1. Redes neurais (Computação). 2. Meteorologia . 3. Radiação solar. I. Título.

---

CDD 658.51

**JEFFERSON LEMOS FERNANDES**

**PREVISÃO DE IRRADIAÇÃO SOLAR ATRAVÉS DE UM MODELO DE REDE  
NEURAL OTIMIZADA COM O ALGORITMO DE BROYDEN-FLETCHER-  
GOLDFARB-SHANNO**

Aprovada em: \_\_\_/\_\_\_/\_\_\_\_\_.

**BANCA EXAMINADORA**

---

Professor Dr. Paulo Alexandre Costa Rocha (Orientador)  
Universidade Federal do Ceará (UFC)

---

Professor Dr. André Valente Bueno  
Universidade Federal do Ceará (UFC)

---

Professor Dr. Daniel Albiero  
Universidade Federal do Ceará (UFC)

À Márcia.

Aos meus pais e aos meus amigos.

## AGRADECIMENTO

Gostaria de agradecer em especial aos meus pais por todo incentivo aos estudos, e à Márcia, por todo apoio e carinho recebidos.

Agradeço ao professor Dr Paulo Alexandre Costa Rocha pelos ensinamentos que me passou durante minha trajetória acadêmica.

A todos os professores e funcionários da UFC.

Aos meus amigos de graduação, Debora Rêgo, Larissa Colares e Welbson Carneiro.

Aos meus queridos amigos, Geovani Rocha, Magno Bezerra, Rayanne Mendes, Rafaela Viana, Raphael Braga, Daniel Cid, Rosangela Pasqualli, Mana Melo, Michael Robson, Lucas Peixoto, Felipe Oliveira, Semirames Albuquerque, Rafael Moraes, Markus Lyma, Caroline Albuquerque, Ingrid Lopes, Little Richard e Vitória Lakiss.

Je tiens à remercier à toute l'équipe de Centrale-Supélec.

Je remercie également à toutes mes chères connasses, Khaoula, Célia, Mathilde, Rosine, Melissa, Claire, Bélanda, Jana et Ella, Nikki, David et Antoine des très beaux moments que vous avez partagés avec moi.

Arigatougozaimasu Mayukochan to Kanokochan ; Watashi wa, anata aishiteimasu.

À Khaoula Soussi, je te remercie énormément de tout ton soutien, ton amitié et tes conneries sans lesquels cette aventure n'aurais certainement pas été aussi formidable.

“Si tu ne saisis pas le petit grain de folie chez quelqu’un, tu ne peux pas l’aimer. Si tu ne saisis pas son point de démente, tu passes à côté. Le point de démente de quelqu’un, c’est la source même de son charme ” (Gilles Deleuze)

## RESUMO

A energia solar é uma fonte de energia renovável bastante promissora no Brasil, onde o sol é abundante e novas instalações de sistemas a base de energia solar recebem cada vez mais incentivo. A fim de avaliar a viabilidade de um sistema solar, é preciso ter conhecimento das condições de irradiação solar incidente. Entretanto, dados históricos de irradiação solar muitas vezes não estão disponíveis com facilidade. O desenvolvimento de modelos de previsão de irradiação solar incidente em uma superfície a partir de variáveis meteorológicas de mais fácil acesso com algoritmos de inteligência artificial representa uma boa alternativa à falta de dados históricos. Diante disto, foram criados modelos de redes neurais para a previsão de irradiação solar global diária, para a média semanal de irradiação solar global e para a média mensal de irradiação solar global. Para isto, foi utilizado o software de programação R e uma série histórica de dados de temperatura máxima, temperatura mínima, umidade relativa, insolação, irradiação extraterrestre, precipitação total, evaporação e nebulosidade correspondente ao período de 1974 a 1988. A partir dos dados diários, obtiveram-se séries com valores médios semanais e mensais. Para cada grupo de dados experimentais, foi possível determinar a quantidade de neurônios na camada oculta da rede neural que minimiza o erro de previsão do modelo ao se analisar o gráfico do erro de previsão dos modelos criados. As redes neurais foram comparadas por meio de seus desempenhos com relação ao erro quadrático médio (RRMSE), erro percentual médio (MAPE) e coeficiente de determinação ( $R^2$ ). Através desta análise, foi demonstrado que o modelo para a previsão da média mensal de irradiação solar global apresenta os melhores resultados.

**Palavras-chave:** Irradiação solar. Redes neurais. Variáveis meteorológicas.

## ABSTRACT

*Solar energy is a promising renewable source of energy in Brazil due to its tropical climate and new solar system installations receive more incentives. In order to evaluate the viability of a solar system, it is important to have access to knowledge about solar radiation conditions throughout the year. However, solar radiation data is not always available. Therefore, the development of solar radiation prediction models based on meteorological variables available is a good alternative to the absence of solar radiation data. In this work a neural network algorithm was developed in order to create and evaluate solar radiation prediction models using meteorological data as inputs. The algorithm is developed using the R free software and maximum temperature, minimum temperature, relative humidity, sunshine hours, extraterrestrial radiation, total precipitation, evaporation, wind speed and cloudiness data from 1974 to 1988. From daily data, a weekly and a monthly series were generated. A neural network was trained on each one of these groups in order to find the best neural network architecture in each case. The neural networks were compared using root mean squared error (RMSE), mean percentage error (MAPE) and coefficient of determination ( $R^2$ ). The analysis shows that the model created for monthly average solar radiation prediction presents the best results.*

**Keywords:** *Solar radiation. Neural networks. Meteorological variables.*

## LISTA DE ILUSTRAÇÕES

Figura 1 – Ângulos solares numa superfície inclinada.....	19
Figura 2 – Diagrama em blocos do sistema nervoso.....	20
Figura 3 – Representação de um neurônio .....	21
Figura 4 – Representação de um perceptron .....	22
Figura 5 – Tipos de tarefas .....	25
Figura 6 – Representação de uma rede neural MLP.....	26
Figura 7 – Estrutura da rede neural modelada.....	38
Figura 8 – Erros de treinamento e teste ideias.....	43
Figura 9 – Erros de teste reais .....	44

## LISTA DE GRÁFICOS

Gráfico 1 – Comportamento dos erros de treinamento e teste a cada iteração	42
Gráfico 2 – Erro versus Complexidade da rede para o grupo diário	45
Gráfico 3 – Erro versus Complexidade da rede para o grupo semanal	45
Gráfico 4 – Erro versus Complexidade da rede para o grupo mensal	46
Gráfico 5 – Erro de treinamento e teste versus Quantidade de dados para treinamento para o grupo diário	47
Gráfico 6 – Erro de treinamento e teste versus Quantidade de dados para treinamento para o grupo semanal	47
Gráfico 7 – Erro de treinamento e teste versus Quantidade de dados para treinamento para o grupo mensal	48
Gráfico 8 – Dados de treinamento, dados de teste de modelo obtido para o grupo mensal	48
Gráfico 9 – Dispersão de treinamento e teste para o grupo mensal	49

## LISTA DE TABELAS

Tabela 1 – Funções de ativação e suas características.....	24
Tabela 2 – Porcentagem de dados experimentais presentes em cada ano .....	39
Tabela 3 – Estatísticas descritivas para o grupo diário.....	41
Tabela 4 – Estatísticas descritivas para o grupo semanal.....	41
Tabela 5 – Estatísticas descritivas para o grupo mensal.....	41
Tabela 5 – Coeficientes de determinação dos modelos .....	49

## SUMÁRIO

<b>1 INTRODUÇÃO</b> .....	14
<b>2 JUSTIFICATIVA</b> .....	16
<b>3 OBJETIVOS</b> .....	17
3.1 Geral .....	17
3.2 Específico .....	17
<b>4 FUNDAMENTAÇÃO TEÓRICA</b> .....	18
4.1 Irradiação solar extraterrestre .....	18
4.2 Introdução às redes neurais .....	20
4.3 <i>Feedforward</i> Perceptron: modelo matemático .....	22
4.4 Função de ativação .....	23
4.5 Problemas de regressão e classificação .....	25
4.6 Modelo de redes neurais de múltiplas camadas (MLP).....	25
4.7 Treinamento da rede neural: A retropropagação do erro.....	28
4.8 Algoritmo de Broyden-Fletcher-Goldfarb-Shanno .....	32
4.9 Avaliação do desempenho de redes neurais .....	33
4.10 Estatística descritiva.....	34
<b>5 METODOLOGIA</b> .....	36
5.1 Desenvolvimento do algoritmo de redes neurais.....	36
5.2 Obtenção de dados experimentais .....	37
5.3 Pré-processamento dos dados experimentais .....	37
5.4 Estudo de caso .....	38
<b>6 RESULTADOS E DISCUSSÃO</b> .....	40
<b>7 CONCLUSÃO</b> .....	51
<b>8 REFERÊNCIAS</b> .....	52

## 1. INTRODUÇÃO

Atualmente, a energia solar esta sendo amplamente aproveitada em várias localidades ao redor do globo a fim de melhorar a sustentabilidade e diminuir os problemas ambientais prevaletentes tais como o aquecimento global e a poluição do ar (SHAMSHIRBAND *et al.*, 2015).

O conhecimento da irradiação solar incidente em um local particular é essencial para aplicações em sistemas agrícolas, hidrológicos, ecológicos bem como aplicações em energia solar. A melhor maneira de se obter dados de irradiação solar global é através de medidas remotas em uma localidade particular usando instrumentos específicos: devido ao alto custo para calibração e manutenção dos instrumentos, dados de irradiação solar são limitados em várias estações meteorológicas ao redor do mundo (HUNT *et al.*, 1998).

As dificuldades e incertezas envolvendo as medidas de irradiação solar global resultaram no desenvolvimento de inúmeros modelos e algoritmos para a sua estimação a partir de algumas variáveis meteorológicas medidas rotineiramente consistindo da insolação, temperatura atmosférica máxima, mínima e média, umidade relativa, nebulosidade, etc. Ao longo dos últimos anos, um vasto número de modelos incluindo modelos empíricos, modelos baseados em satélites e algoritmos estocásticos têm sido desenvolvidos para estimar a irradiação global incidente em uma superfície horizontal (OLATOMIWA *et al.*, 2015).

Recentemente, inteligência artificial e técnicas de inteligência computacional são extensivamente utilizadas a fim de resolver problemas do mundo real onde metodologias convencionais são inadequadas ou uma precisão adicional é requerida. Aplicações de tais abordagens no campo de estimação de irradiação solar têm recebido atenção específica nos últimos anos (SHAMSHIRBAND *et al.*, 2015).

Bosch *et al.* (2008) conduziram uma investigação a fim de estimar a irradiação solar global diária usando redes neurais artificiais em uma região montanhosa situada ao sudeste da espanha. Seus resultados demonstraram que redes neurais artificiais podem ser consideradas como uma técnica efetiva de fácil manipulação para a estimativa da irradiação solar em terrenos montanhosos. Benghanem *et al.* (2009) desenvolveram seis modelos baseados em redes neurais para estimar a irradiação global incidente horizontalmente em Madinah, na Arábia Saudita. Eles utilizaram diferentes conjuntos de variáveis para cada

modelo e seus resultados evidenciaram uma maior precisão dos modelos que tinham insolação e temperatura do ar.

Percebe-se que a utilização de redes neurais se mostra uma alternativa razoável aos diversos tipos de modelos existentes.

Dito isto, o presente trabalho tem como consistiu em desenvolver um modelo de previsão de irradiação solar global a partir de parâmetros meteorológicos como insolação, temperatura do ar, umidade relativa, etc. Para isto, foi desenvolvido um algoritmo utilizando a linguagem de programação R.

## 2. JUSTIFICATIVA

O Brasil possui um grande potencial de exploração de energia solar, com previsão de que até 2024, 700 mil consumidores residenciais e comerciais terão instalado em seus telhados painéis solares, segundo o Ministério de Minas e Energia (MME).

A irradiação solar global incidente é um parâmetro de grande importância para a viabilização de projetos e aplicações envolvendo a geração de energia solar. Entretanto, muitos sítios não possuem estações meteorológicas ou uma base de dados de valores medidos de irradiação solar que possibilitem uma melhor tomada de decisão na elaboração de projetos.

Segundo Kathib *et al.* (2012), o uso da energia solar é especialmente importante num contexto de altos preços dos combustíveis fósseis e da degradação do meio ambiente. Entretanto, segundo Shamshirband *et al.* (2015), a disponibilidade de dados de irradiação solar é um requisito fundamental para que os especialistas em sistemas solares simulem, operem e monitorem com sucesso as tecnologias de energia solar em suas diversas aplicações.

Diante da importância deste parâmetro, este trabalho consiste em desenvolver um modelo de rede neural a ser utilizado na previsão da irradiação solar a partir de variáveis meteorológicas de mais fácil acesso, aperfeiçoar este modelo através do algoritmo de BFGS e testá-lo em dados experimentais a fim de discutir o desempenho obtido.

### **3. OBJETIVOS**

#### **3.1. Geral**

O presente trabalho teve como objetivo geral o desenvolvimento de um modelo de rede neural a fim realizar a previsão da irradiação solar global incidente a partir de parâmetros meteorológicos. Para isso, foi utilizada a linguagem de programação R.

#### **3.2. Específico**

Para alcançar o objetivo geral, as seguintes etapas foram necessárias:

- a) Desenvolvimento do algoritmo de rede neural;
- b) Otimização deste algoritmo através do algoritmo de Broyden-Fletcher-Goldfarb-Shanno (BFGS);
- c) Obtenção dos dados experimentais;
- d) Pré-processamento dos dados experimentais;
- e) Aplicação destes algoritmos no estudo da previsão de irradiação solar.

## 4. FUNDAMENTAÇÃO TEÓRICA

### 4.1. Irradiação solar extraterrestre

O sol é a estrela mais importante do sistema solar e sua estrutura e características são responsáveis pela energia que ele irradia no espaço em direção a Terra.

A irradiação solar é a principal fonte de energia na maioria das interações entre a atmosfera e a superfície e é responsável por quase todos os processos físicos, químicos e biológicos na atmosfera terrestre e nos sistemas oceânicos. O conhecimento das componentes global, difusa e direta da irradiação solar na superfície é crucial para determinar áreas com potencial para a geração de energia solar, para avaliar o consumo energético de construções, para subsidiar estudos ecofisiológicos, para estimar a evapotranspiração de cereais e para apoiar o planejamento urbano.

Além disto, devido às razões ambientais e econômicas, países tropicais como o Brasil, precisarão utilizar seus potenciais de energia solar, estimado entre 1500 e 2200  $kW h m^{-2}$  ao ano, para diversificar as suas matrizes energéticas. Atualmente, o Brasil possui uma capacidade instalada de 134 GW para geração de energia, das quais 5% são provenientes da energia solar e da energia eólica (MARQUES FILHO *et al.*, 2016).

Por consequência, é importante entender como a irradiação solar chega a superfície terrestre e o que influencia a magnitude desta energia.

Segundo Duffie *et al.* (2013), a irradiação emitida pelo sol e sua posição em relação a terra são responsáveis pela intensidade de irradiação solar uniforme fora da atmosfera terrestre. Esta contante, chamada de constante solar,  $G_{CS}$ , é a energia emitida pelo sol por unidade de tempo que é recebida em uma unidade de área da superfície perpendicular à direção de propagação da irradiação no ponto médio da distância entre o sol e a terra da superfície externa da atmosfera terrestre. O valor da constante solar é dado pela equação 1.

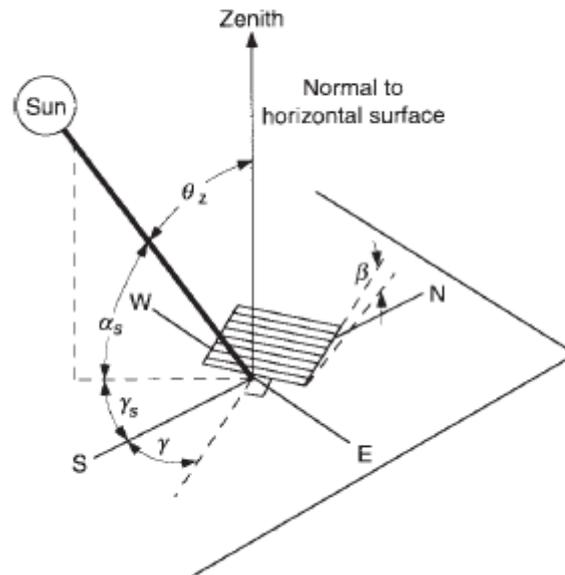
$$G_{CS} = 1367 W/m^2 \quad [1]$$

A irradiação incidente em um determinado ponto da superfície terrestre é influenciada pela localização deste ponto em relação ao sol.

As relações geométricas entre um plano com uma orientação particular em relação à terra e em um determinado momento (este plano sendo fixo ou se movendo em relação a terra) e a irradiação solar direta, isto é, a posição do sol em relação ao plano, pode ser descrita em termo de vários ângulos (DUFFIE *et al.*, 2013).

A Figura 1 mostra alguns destes ângulos.

Figura 1 – Ângulos solares numa superfície inclinada



Fonte: DUFFIE *et al.*, 2013.

Segundo Duffie *et al.* (2013), a irradiação solar extraterrestre pode ser calculada por meio da equação 2:

$$H_0 = \frac{24 \cdot 3600 \cdot G_{cs}}{\pi} * \left(1 + 0,033 * \cos\left(\frac{360n}{365}\right)\right) * (\cos\varphi \cos\delta \cos\omega_s + \frac{\pi\omega_s}{180} * \text{sen}\varphi \text{sen}\delta) \quad [2]$$

Onde,  $n$  é o número do dia,  $\varphi$  é a latitude,  $G_{cs}$  é a constante solar,  $\delta$  é o ângulo de declinação e  $\omega_s$  é o ângulo horário de pôr do sol.

$$\delta = 23,45 * \text{sen}\left(360 * \frac{284+n}{365}\right) \quad [3]$$

$$\omega_s = -\tan\varphi \tan\delta \quad [4]$$

$$-90^\circ \leq \varphi \leq 90^\circ \quad [5]$$

$$1 \leq n \leq 365 \quad [6]$$

A irradiação solar normal à superfície da terra é sofre variações devido às mudanças na irradiação extraterrestre e dois outros fenômenos significativos: espalhamento atmosférico causado por moléculas de ar, água e poeira, e absorção atmosférica causada pela presença de  $O_2$ ,  $H_2O$  e  $CO_2$  (DUFFIE *et al.*, 2013).

Trabalhos extensivos sobre modelos empíricos têm sido feitos a fim de se estimar a irradiação solar global em uma superfície horizontal utilizando medidas de insolação. Vários modelos convencionais de previsão de irradiação solar global a partir de variáveis

meteorológicas foram apresentados pelos pesquisadores. Utilizar redes neurais artificiais em modelos de previsão onde as variáveis não possuem nenhuma relação específica provou sua eficiência. Variáveis meteorológicas e climatológicas são os fatores mais compreensivos e importantes na indicação da quantidade de irradiação solar em um sítio. Vários estudos utilizando redes neurais artificiais para a previsão de irradiação solar têm sido apresentados (BEHRANG *et al.*, 2010).

Portanto, é sabido que as condições meteorológicas na atmosfera terrestre influenciam a quantidade de irradiação solar incidente em uma superfície.

#### 4.2. Introdução às redes neurais

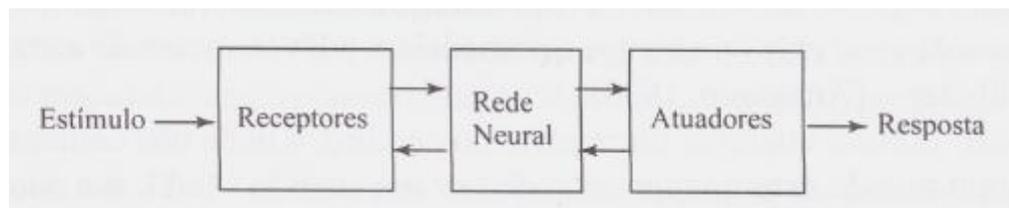
As Redes Neurais, ou redes artificiais, representam uma tecnologia que tem raízes em muitas disciplinas: neurociência, matemática, estatística, física, ciência da computação e engenharia. As redes neurais encontram aplicações em campos tão diversos, como modelagem, análise de séries temporais, reconhecimento de padrões, processamento de sinais e controle, em virtude de uma importante propriedade: a habilidade de aprender a partir de dados de entrada com ou sem professor (HAYKIN, 2008).

Segundo Bishop (2007), o termo rede neural tem suas origens nas tentativas de encontrar uma representação matemática de como o processamento de informações acontece em sistemas biológicos.

Desta forma, o funcionamento das redes neurais é baseado no funcionamento dos neurônios encontrados em um cérebro humano.

Segundo Haykin (2008), o sistema nervoso pode ser visto como um sistema de três estágios, como mostrado no diagrama em blocos da Figura 2.

Figura 2 – Diagrama em blocos do sistema nervoso

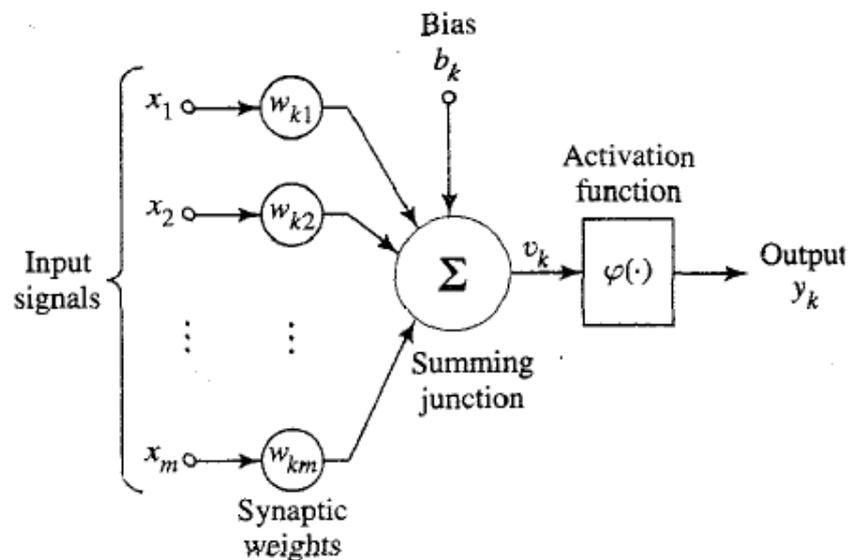


Fonte: HAYKIN, 2001.

A partir da Figura 2, percebe-se que o sistema nervoso funciona em dois sentidos. No sentido da esquerda para a direita, estímulos do corpo humano e do ambiente externo são recebidos pelos receptores cuja função é de transmití-los à rede neural para que esta proceda a tomadas de decisões, gerando uma resposta aos sinais recebidos. No sentido da direita para a esquerda, as respostas geradas pela rede neural são reintroduzidas no sistema, chegando até os receptores para que o sistema como um todo esteja mais adaptado quando um mesmo estímulo é repetido. Em outras palavras, o sistema nervoso utiliza experiências passadas a fim de tomar decisões a respeito de estímulos futuros.

A Figura 3 mostra a representação de um neurônio para aplicações em redes neurais artificiais.

Figura 3 – Representação de um neurônio



Fonte: HAYKIN, 2001.

Os principais componentes de um neurônio são os sinais de entrada, os pesos sinápticos, o *bias* aplicado, a junção aditiva e a função de ativação deste neurônio.

Os sinais de entrada representam os estímulos recebidos pelo neurônio e que se transformarão em um sinal de saída por meio dos diferentes processos sofridos ao longo do neurônio.

Os pesos sinápticos representam um elo de conexão entre os sinais de entrada e o neurônio em questão. Cada sinal de entrada é conectado ao neurônio por um elo de conexão diferente.

O termo bias tem o efeito de aumentar ou diminuir a informação total que é levada em consideração pelo neurônio a fim de gerar uma resposta ao estímulo inicial.

A junção aditiva, normalmente um combinador linear, capta a informação dos diferentes sinais de entrada ponderados pelos seus pesos sinápticos e do termo bias, gerando uma informação líquida a ser tratada pelo neurônio.

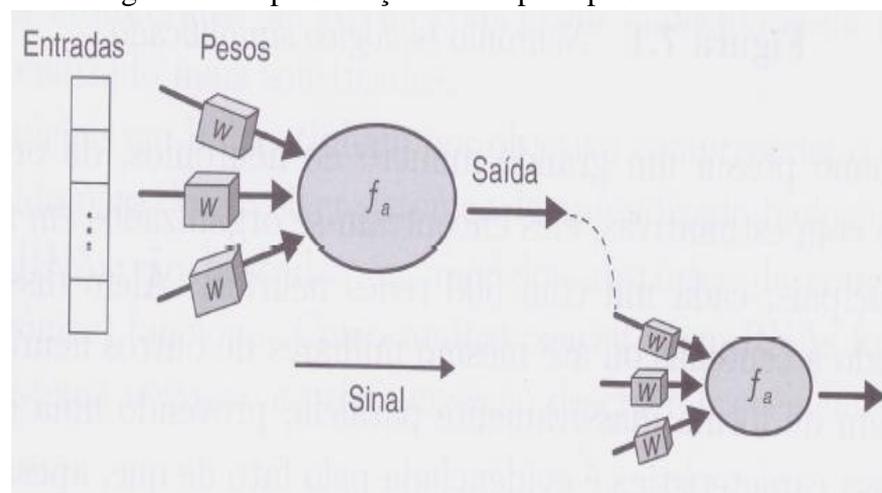
Finalmente, a função de ativação de um neurônio é responsável por transformar a informação líquida recebida pelo neurônio em uma informação de saída, ou seja, ela fornece uma resposta aos estímulos recebidos pelo neurônio em questão.

#### 4.3. *Feedforward Perceptron*: modelo matemático

Segundo Haykin (2008), o perceptron é a forma mais simples de uma rede neural usada para a classificação de padrões. Ele consiste de um único neurônio com pesos sinápticos e bias.

A Figura 4 mostra um exemplo de um perceptron a partir do qual as equações a seguir foram originadas.

Figura 4 – Representação de um perceptron



Fonte: FACELI *et al.*, 2011.

Os perceptrons recebem sinais de entradas representados pelo vetor coluna  $x$  dado pela equação [7]:

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} \quad [7]$$

A cada uma destas entradas, é atribuído um peso sináptico  $w_{ij}$ , onde  $j$  representa o sinal de entrada  $x_j$  e  $i$  representa o neurônio analisado.

Desta forma, pode-se representar o vetor dos pesos sinápticos pela equação [8]:

$$w_i = [w_{i1} \ w_{i2} \ w_{i3} \ \dots \ w_{in}] \quad [8]$$

Em seguida, a junção aditiva é aplicada aos sinais de entrada e seus respectivos pesos sinápticos. Esta junção aditiva é uma combinação linear e pode ser dada pela equação [9]:

$$a_i = w_{i1}x_1 + w_{i2}x_2 + \dots + w_{in}x_n + b_i = w_i * x_i + b_i \quad [9]$$

Onde  $b_i$  indica o termo bias aplicado ao neurônio  $i$ .

As quantidades  $a_i$  são conhecidas como ativações. Cada uma delas é transformada utilizando uma função de ativação  $\varphi(\cdot)$  para gerar o sinal de saída do neurônio como mostrado pela equação [10]:

$$z_i = \varphi(z_i) \quad [10]$$

Ao processo desde a entrada dos sinais  $x$  até a geração do sinal de saída por meio da utilização da função de ativação, é dado o nome de *feedforward propagation*, ou propagação em frente.

#### 4.4. Função de ativação

Em redes neurais artificiais, a função de ativação  $\varphi(\cdot)$  de um neurônio representa a taxa com a qual determinado neurônio está sendo ativado ou utilizado.

Cada neurônio em uma rede neural recebe uma série de entradas. Uma função chamada de função de ativação é aplicada a esses valores de entrada, o que resulta no nível de ativação do neurônio, que é o valor de saída do neurônio. Há uma série de funções que podem ser utilizadas nos neurônios (COPPIN, 2004).

As funções de ativação têm um papel fundamental no treinamento de redes neurais e é importante que elas possuam algumas características a fim de melhorar o desempenho do treinamento das redes neurais.

Segundo HAYKIN (2008), a função de ativação deve ser diferenciável para que se possa utilizar um algoritmo que visa a minimizar o erro das tarefas realizadas pelas redes neurais.

Além disto, se a função de ativação for diferenciável mais de uma vez, é possível desenvolver algoritmos que extraíam informações de segunda ordem acerca do erro cometido pelas redes neurais, o que possibilita um treinamento mais rápido e eficiente.

As funções de ativação mais comuns e suas derivadas estão listadas na Tabela 1.

Tabela 1 – Funções de ativação e suas características

Função de ativação	Equação	Derivada	Intervalo de valores	Ordem de continuidade
Passo binário	$\varphi(x) = \begin{cases} 0, & x \leq 0 \\ 1, & x > 0 \end{cases}$	$\varphi(x) = 0, \quad x \neq 0$	[0,1]	$C^{-1}$
Identidade	$\varphi(x) = x$	$\varphi'(x) = 1$	$]-\infty, +\infty[$	$C^\infty$
Sigmoide	$\varphi(x) = \frac{1}{1 + e^{-x}}$	$\varphi'(x) = \varphi(x) * (1 - \varphi(x))$	[0,1]	$C^\infty$
Softplus	$\varphi(x) = \ln(1 + e^x)$	$\varphi'(x) = 1 - \varphi(x)^2$	$[0, +\infty[$	$C^\infty$
Tangente hiperbólica	$\varphi(x) = \tanh(x)$	$\varphi'(x) = \frac{1}{1 + e^{-x}}$	$[-1,1]$	$C^\infty$

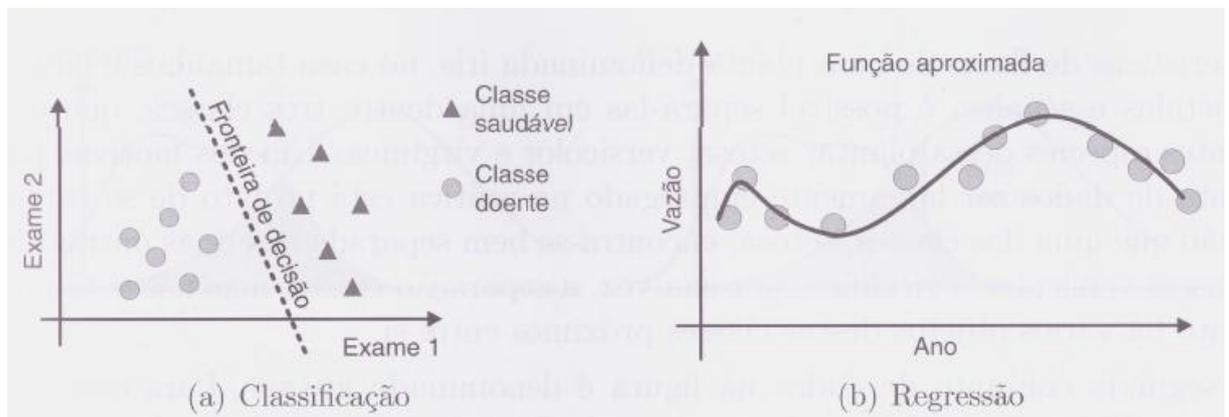
Fonte: Elaborada pelo autor.

#### 4.5. Problemas de regressão e classificação

Segundo Feceli *et al.* (2011), um algoritmo de aprendizado de máquina preditivo é uma função que, dado um conjunto de exemplos rotulados, constrói um estimador. O rótulo ou etiqueta toma valores num domínio conhecido. Se esse domínio for um conjunto de valores nominais, tem-se um problema de classificação, também conhecido como aprendizado de conceitos, e o estimador gerado é um classificador. Se o domínio for um conjunto infinito e ordenado de valores, tem-se um problema de regressão, que induz um regressor.

A Figura 5 ilustra os dois tipos de problemas.

Figura 5 – Tipos de tarefas



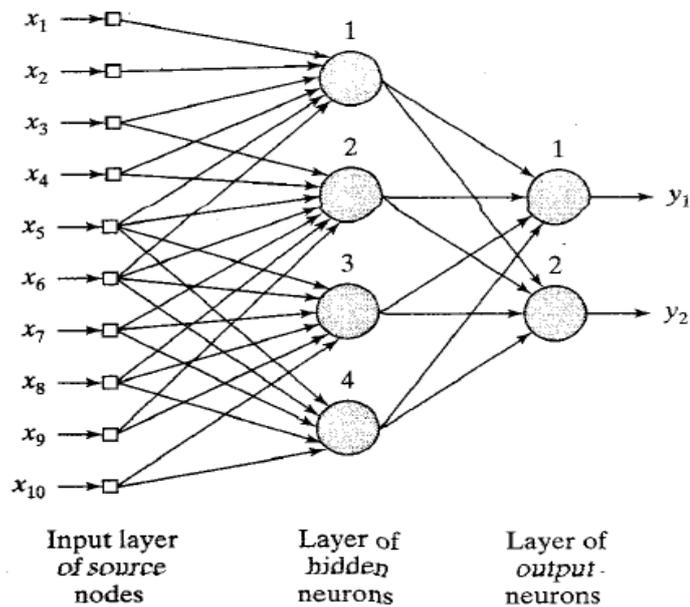
Fonte: FACELI *et al.*, 2011.

Em outras palavras, é possível realizar os dois tipos de tarefas por meio de redes neurais. Ao mostrar exemplos cujo resultado é conhecido, a rede neural aprende a prever um valor ou uma classificação quando encontra um novo conjunto de entradas.

#### 4.6. Modelo de redes neurais de múltiplas camadas (MLP)

Um modelo de redes neurais de múltiplas camadas, ou *multi layer perceptron* (MLP), é formado a partir da união de vários neurônios, ligados entre si por meio de elos sinápticos e bias. A Figura 6 mostra um exemplo de uma rede neural a partir da qual serão abordados os conceitos abordados na seção anterior.

Figura 6 – Representação de uma rede neural MLP



Fonte: HAYKIN, 2001.

O fluxo de sinal percorre a rede neural da esquerda para a direita através do mecanismo de *forward propagation* visto anteriormente.

Deste modo, pode-se realizar a modelização matemática da rede neural como uma cadeia de perceptrons conectados uns aos outros, onde o sinal de saída de uma camada anterior representa o sinal de entrada da camada seguinte até que o sinal resultante saia da última camada de neurônios.

Para a camada de entrada, com D diferentes sinais, tem-se o vetor de entrada dado pela equação [11]:

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_D \end{bmatrix} \quad [11]$$

Da camada de entrada à primeira camada oculta, pode-se definir uma matriz de pesos sinápticos dada pela equação [12]:

$$w_{ij}^{(1)} = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1D} \\ w_{21} & w_{22} & \cdots & w_{2D} \\ \vdots & \ddots & & \vdots \\ w_{M1} & w_{M2} & & w_{MD} \end{bmatrix} \quad [12]$$

Onde  $w_{ij}^{(1)}$  é o peso sináptico do sinal de entrada  $j$  no neurônio de entrada  $i$  e o superescrito (1) indica que os parâmetros correspondentes estão na primeira camada da rede neural.

Da mesma forma, para cada camada, define-se um vetor coluna  $[b]_{M \times 1}^n$ , onde  $n$  é o número da camada e  $M$ , a quantidade de neurônios nesta camada.

Com isto, a ativação no neurônio  $i$  da camada  $n$  pode ser obtida por meio da equação [13]:

$$a_i^{(n)} = w_{i1}^{(n)} x_1 + w_{i2}^{(n)} x_2 + \cdots + w_{iD}^{(n)} x_D + [bias]_i^{(n)} \quad [13]$$

Para todos os neurônios da camada, pode-se escrever a notação acima como uma matriz de ativações, dada pela equação [14]:

$$z_i^{(n)} = \begin{bmatrix} w_{11}^{(n)} & w_{12}^{(n)} & \cdots & w_{1D}^{(n)} \\ w_{21}^{(n)} & w_{22}^{(n)} & \cdots & w_{2D}^{(n)} \\ \vdots & \ddots & & \vdots \\ w_{i1}^{(n)} & w_{i2}^{(n)} & & w_{iD}^{(n)} \end{bmatrix} * \begin{bmatrix} x_1^{(n)} \\ x_2^{(n)} \\ \vdots \\ x_D^{(n)} \end{bmatrix} + \begin{bmatrix} bias_1^{(n)} \\ bias_2^{(n)} \\ \vdots \\ bias_i^{(n)} \end{bmatrix} \quad [14]$$

$$z_i^{(n)} = \begin{bmatrix} \sum_{j=1}^D w_{1j}^{(n)} * x_j^{(n)} + b_1^{(n)} \\ \sum_{j=1}^D w_{2j}^{(n)} * x_j^{(n)} + b_2^{(n)} \\ \vdots \\ \sum_{j=1}^D w_{ij}^{(n)} * x_j^{(n)} + b_i^{(n)} \end{bmatrix} \quad [15]$$

Em seguida, podem-se calcular os sinais de saída de cada neurônio por meio da aplicação da função de ativação como mostrado pelas equações [16] e [17]:

$$a_i^{(n)} = \varphi \left( \begin{bmatrix} w_{11}^{(n)} & w_{12}^{(n)} & \cdots & w_{1D}^{(n)} \\ w_{21}^{(n)} & w_{22}^{(n)} & \cdots & w_{2D}^{(n)} \\ \vdots & \ddots & & \vdots \\ w_{M1}^{(n)} & w_{M2}^{(n)} & & w_{MD}^{(n)} \end{bmatrix} * \begin{bmatrix} x_1^{(n)} \\ x_2^{(n)} \\ \vdots \\ x_D^{(n)} \end{bmatrix} + \begin{bmatrix} b_1^{(n)} \\ b_2^{(n)} \\ \vdots \\ b_i^{(n)} \end{bmatrix} \right) \quad [16]$$

$$a_i^{(n)} = \begin{bmatrix} \varphi\left(\sum_{j=1}^D w_{1j}^{(n)} * x_j^{(n)} + b_1^{(n)}\right) \\ \varphi\left(\sum_{j=1}^D w_{2j}^{(n)} * x_j^{(n)} + b_2^{(n)}\right) \\ \vdots \\ \varphi\left(\sum_{j=1}^D w_{Mj}^{(n)} * x_j^{(n)} + b_i^{(n)}\right) \end{bmatrix} \quad [17]$$

Com isto, para o exemplo da Figura 4, onde existe uma camada de entrada, uma camada oculta e uma camada de saída, o resultado final na camada de saída é dado pela equação [18]:

$$y_1(\mathbf{x}, \mathbf{w}) = a_1^{(2)} \quad [18]$$

$$y_1(\mathbf{x}, \mathbf{w}) = h\left(\sum_{k=1}^M w_{1k}^{(2)} * \varphi\left(\sum_{j=1}^D w_{1j}^{(1)} * x_j^{(1)} + b_1^{(1)}\right) + b_1^{(2)}\right) \quad [19]$$

Onde a notação  $y_1(\mathbf{x}, \mathbf{w})$  indica o resultado final do neurônio 1 da última camada, ou camada de saída, com relação ao vetor de sinais de entrada  $\mathbf{x}$  e os pesos sinápticos da rede,  $\mathbf{w}$ .

Por consequência, uma rede neural de múltiplas camadas é definida pelo seu número de camadas, pela quantidade de neurônios em cada camada, pelas funções de ativação utilizadas em cada neurônio e por seus pesos sinápticos e bias. Com estas informações, é possível obter um sinal de saída para cada sinal de entrada.

Segundo Haykin (2008), os perceptrons de múltiplas camadas têm sido aplicados com sucesso para resolver problemas difíceis, através do seu treinamento de forma supervisionada com um algoritmo muito popular conhecido como algoritmo de retropropagação de erro (*error backpropagation*).

#### 4.7. Treinamento da rede neural: A retropropagação do erro

O desenvolvimento do algoritmo de retropropagação representa um marco nas redes neurais, pois fornece um método computacional eficiente para o treinamento de perceptrons de múltiplas camadas. Apesar de não podermos afirmar que o algoritmo de retropropagação forneça uma solução ótima para todos os problemas resolúveis, ele acabou com o pessimismo sobre a aprendizagem em máquinas de múltiplas camadas (HAYKIN, 2008).

O algoritmo de retropropagação do erro visa a reajustar os pesos sinápticos e bias das conexões dentro da rede neural por meio de uma avaliação do erro obtido quando se apresenta um conjunto de sinais de entrada com sinais de saída conhecidos.

Desta forma, é possível realizar um treinamento de uma rede neural num conjunto de dados disponíveis para que a mesma possa ser utilizada na previsão de sinais de saída desconhecidos.

Assumindo um conjunto de  $N$  exemplos constituídos de sinais de entrada  $\mathbf{x}$  e seus respectivos sinais de saída  $\mathbf{y}$ , pode-se definir o erro de um sinal de saída como na equação [20]:

$$e(n) = \hat{y}(n) - y(n) \quad [20]$$

Define-se a função custo quadrático médio para um neurônio como na equação [21]:

$$J(w, bias) = \frac{1}{2N} \sum_1^N e^2(n) \quad [21]$$

A função custo representa uma medida do desempenho do treinamento da rede neural. O objetivo do algoritmo de retropropagação do erro é encontrar os parâmetros  $\tilde{w}$  e  $\tilde{b}$  que minimizam a função custo.

O algoritmo da retropropagação visa a encontrar os valores de  $\tilde{w}$  e  $\tilde{b}$  para os quais a função custo assume um valor mínimo. Uma das condições para que um par  $(\tilde{w}, \tilde{b})$  indique um mínimo da função custo é a de que o gradiente em torno deste ponto seja aproximadamente zero.

O gradiente da função custo pode ser obtido por meio da equação [22]:

$$\frac{\partial J(w,b)}{\partial(w,b)} = \frac{\partial \frac{1}{2N} \sum_1^N e^2(n)}{\partial(w,b)} \quad [22]$$

$$\frac{\partial J(w,b)}{\partial(w,b)} = \frac{1}{2N} \frac{\partial \sum_1^N e^2(n)}{\partial(w,b)} \quad [23]$$

A derivada da soma equivale à soma das derivadas e pode-se reescrever [23] por meio da equação [24]:

$$\frac{\partial J(w,b)}{\partial(w,b)} = \frac{1}{2N} * \sum_1^N \frac{\partial e(n)^2}{\partial(w,b)} \quad [24]$$

$$\frac{\partial J(w,b)}{\partial(w,b)} = \frac{1}{2N} * 2 * \sum_1^N e(n) \frac{\partial e(n)}{\partial(w,b)} \quad [25]$$

$$\frac{\partial J(w,b)}{\partial(w,b)} = \frac{1}{N} * \sum_1^N e(n) \frac{\partial(\hat{y}(n) - y(n))}{\partial(w,b)} \quad [26]$$

$$\frac{\partial J(w,b)}{\partial(w,b)} = \frac{1}{N} * \sum_1^N e(n) \left( \frac{\partial \hat{y}(n)}{\partial(w,b)} - \frac{\partial y(n)}{\partial(w,b)} \right) \quad [27]$$

$$\frac{\partial J(w,b)}{\partial(w,b)} = \frac{1}{N} * \sum_1^N e(n) \left( \frac{\partial \hat{y}(n)}{\partial(w,b)} \right) \quad [28]$$

Sabe-se que,

$$\hat{y}(x, w) = a = \varphi(z) = \varphi(w * x + b) \quad [29]$$

$$\hat{y}(x, w) = \varphi(z) = \varphi(w * x + b) \quad [30]$$

$$\hat{y}(x, w) = \varphi(w * x + b) \quad [31]$$

Logo,

$$\frac{\partial \hat{y}(n)}{\partial(w,b)} = \frac{\partial \varphi(w*x+b)}{\partial(w,b)} = x * \varphi'(w * a + b) = x * \varphi'(z) \quad [32]$$

$$\frac{\partial \hat{y}(n)}{\partial(w,b)} = x * \varphi'(w * a + b) = x * \varphi'(z) \quad [33]$$

$$\frac{\partial \hat{y}(n)}{\partial(w,b)} = x * \varphi'(z) \quad [34]$$

De onde se conclui que,

$$\frac{\partial J(w,b)}{\partial(w,b)} = \frac{1}{N} * \sum_1^N e(n) * \varphi'(z) * x_i \quad [35]$$

A correção aplicada aos pesos sinápticos e bias é definida pela regra delta, dada pela equação [36]:

$$\Delta w_{ji}(n) = -\gamma * \frac{\partial J(w,b)}{\partial(w,b)} \quad [36]$$

Ou seja, a correção é feita no sentido de descréscimo da função de custo a fim de se atingir o mínimo.

O termo  $\gamma$  é o parâmetro da taxa de aprendizagem do algoritmo de retropropagação. Ele é responsável por determinar a magnitude da mudança dos pesos sinápticos no sentido de descida do gradiente.

O uso da equação [36] em [35], resulta em:

$$\Delta w_{ji}(n) = -\gamma * \frac{1}{N} \sum_1^N e(n) * \varphi'(z) * x_i \quad [37]$$

Define-se o gradiente local por:

$$\delta_j(n) = -\frac{1}{N} \sum_1^N e(n) * \varphi'(z) \quad [38]$$

Logo:

$$\Delta w_{ji}(n) = \gamma * \delta_j(n) * x_i \quad [39]$$

Desta forma, o gradiente local aponta no sentido das modificações necessárias nos pesos sinápticos.

Além disto, existem dois casos a serem analisados: o caso de um neurônio situado na camada de saída e o caso de um neurônio situado em uma camada oculta.

Caso o neurônio se encontre na última camada, utiliza-se a equação [20] a fim de calcular o erro e a equação [38] a fim de calcular o gradiente local diretamente.

Caso o neurônio se encontre em uma camada oculta, o gradiente local de um neurônio  $j$  pertencente à camada oculta é dado por:

$$\delta_j(n) = -\frac{\partial J(w,b)}{\partial z_j} \frac{\partial z_j}{\partial a_j} \quad [40]$$

$$\delta_j(n) = -\frac{\partial J(w,b)}{\partial z_j} \varphi'(z) \quad [41]$$

Este neurônio na camada oculta está ligado ao neurônio  $k$  pertencente à camada de saída. Desta forma, o erro deste neurônio é dado por:

$$J(w, bias) = \frac{1}{2N} \sum_1^N e_k^2(n) \quad [42]$$

Diferenciando esta função em relação ao sinal de saída do neurônio  $j$ , encontra-se:

$$\frac{\partial J(w,b)}{\partial a_j} = \frac{1}{N} \sum_1^N e_k(n) * \frac{\partial e_k}{\partial a_j} \quad [43]$$

Utilizando a regra da cadeia, tem-se:

$$\frac{\partial J(w,b)}{\partial a_j} = \frac{1}{N} * \sum_1^N e_k(n) * \frac{\partial e_k}{\partial z_k} \frac{\partial z_k}{\partial a_j} \quad [44]$$

$$\frac{\partial e_k}{\partial z_k} = \frac{\partial (\widehat{y}_k - y_k)}{\partial z_k} \quad [45]$$

$$\frac{\partial e_k}{\partial z_k} = \frac{\partial \varphi(z_k)}{\partial z_k} \quad [46]$$

$$\frac{\partial e_k}{\partial z_k} = \varphi'(z_k) \quad [47]$$

$$\frac{\partial z_k}{\partial a_j} = w_{kj} \quad [48]$$

$$\text{Pois } z_k = \sum_{j=1}^M w_{kj} * a_j$$

Logo, a equação [44], se transforma em:

$$\frac{\partial J(w,b)}{\partial z_j} = \frac{1}{N} * \sum_1^N e_k(n) * \varphi'(z_k) w_{kj}(n) \quad [49]$$

E podemos escrever a equação [49] da seguinte forma:

$$\frac{\partial J(w,b)}{\partial z_j} = \frac{1}{N} * \sum_1^N \delta_k(n) w_{kj}(n) \quad [50]$$

Onde,

$$\delta_j(n) = -\varphi'(z_j)(n) \sum_1^N \delta_k(n) w_{kj}(n) \quad [51]$$

Com isto, a correção nos pesos é definida pela regra geral:

$$\begin{pmatrix} \text{Correção} \\ \text{de peso} \\ \Delta w_{ji}(n) \end{pmatrix} = \begin{pmatrix} \text{Taxa de} \\ \text{aprendizagem} \\ \gamma \end{pmatrix} * \begin{pmatrix} \text{Gradiente} \\ \text{local} \\ \delta_j(n) \end{pmatrix} * \begin{pmatrix} \text{Sinal de entrada} \\ \text{do neurônio } j \\ x_i(n) \end{pmatrix} \quad [52]$$

Onde:

$$\begin{aligned} & \text{Gradiente} \\ & \text{local} = \\ & \delta_j(n) \\ & \begin{cases} -\frac{1}{N} \sum_1^N e(n) * \varphi'(z), & \text{se } j \text{ é um neurônio da última camada} \\ -\frac{1}{N} \varphi'(z_j)(n) \sum_1^N \delta_k(n) w_{kj}(n), & \text{se } j \text{ é um neurônio da camada oculta} \end{cases} \end{aligned} \quad [53]$$

#### 4.8. Algoritmo de Broyden-Fletcher-Goldfarb-Shanno

Segundo Chun-Cheng Peng *et al.* (2001), no contexto de otimização sem restrições, métodos quase-Newton são algoritmos bastante conhecidos para encontrar um mínimo local de funções como a função custo de redes neurais. Os métodos quase-Newton são baseados no método de Newton para encontrar um ponto estacionário de uma função, onde o gradiente é zero.

O objetivo principal do algoritmo BFGS é encontrar uma direção de descida e um passo de descida que façam com que o aprendizado termine mais rápido. Por consequência, este algoritmo utiliza a informação da derivada de segunda ordem da função custo. Esta informação é representada pela aproximação da matriz Hessiana,  $B$ .

Segundo Yu-Hong Daí (2002), o algoritmo pode ser resumido da seguinte forma:

Passo 0. Dados  $x_1 \in \mathfrak{R}^n, B_1 \in \mathfrak{R}^{n \times n}$  definido positivo

Calcular  $g_1 = \nabla f(x_1)$ . Se  $g_1 = 0$ , parar; se não, escolher  $k = 1$ .

Passo 1. Escolher  $d_k = -B_k^{-1}g_k$ .

Passo 2. Realizar uma pesquisa linear ao longo da direção  $d_k$ , obtendo um valor de  $\gamma_k > 0$ ,  $x_{k+1} = x_k + \gamma_k * d_k$ , e  $g_{k+1} = \nabla f(x_{k+1})$ ;

Se  $g_{k+1} = 0$ , parar.

Passo 3. Escolher

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{\gamma_k \gamma_k^T}{s_k^T \gamma_k} \quad [54]$$

Onde,

$$s_k = \gamma_k d_k \quad [55]$$

$$y_k = g_{k+1} - g_k \quad [56]$$

Passo 4.  $k := K + 1$ ; Ir ao Passo 1.

Ainda segundo Yu-Hong Daí (2002), no algoritmo de BFGS, é necessário que a taxa de aprendizado  $\alpha_k$  obedeça às condições de Wolfe dadas a seguir:

$$f(x_k + \gamma_k d_k) - f(x_k) \leq \delta_1 \gamma_k d_k^T g_k \quad [57]$$

$$d_k^T \nabla f(x_k + \gamma_k d_k) \geq \delta_2 d_k^T g_k \quad [58]$$

Neste algoritmo,  $f$  é a função a ser minimizada, ou seja, a função custo durante o treinamento da rede neural. O parâmetro  $x_k$  representa os pesos e bias durante a iteração  $k$ .

#### 4.9. Avaliação do desempenho de redes neurais

A qualidade dos modelos desenvolvidos a partir de algoritmos de redes neurais pode ser avaliada por meio de parâmetros de erro, como o erro quadrático médio (MSE) e o erro percentual médio (MAPE); bem como por meio do coeficiente de determinação  $R^2$ .

Intuitivamente, o erro quadrático médio é o parâmetro de custo que tende a diminuir a cada iteração até chegar a um ponto de mínimo durante o treinamento de uma rede neural.

O erro percentual médio é uma medida do desempenho de um modelo de previsão que pode ser visto como a distância percentual entre o que realmente ocorreu e o que foi previsto pelo modelo.

Pode-se também utilizar o coeficiente de determinação  $R^2$  como estatística global para avaliar o ajuste do modelo.

Segundo Doreswamy *et al.* (2013), o coeficiente de determinação mede a variância que é interpretada pelo modelo, que significa a redução na variância ao se utilizar o modelo. Tal coeficiente varia de 0 a 1 e o modelo possui uma boa capacidade preditiva quando o  $R^2$  está próximo a 1, assim como o modelo não possui boa capacidade preditiva quando o  $R^2$  está próximo a 0.

O cálculo de todos estes parâmetros se encontra abaixo:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2} \quad [59]$$

$$RRMSE = \frac{RMSE}{\frac{1}{N} \sum_{i=1}^N y_i} \quad [60]$$

$$MAPE = \frac{1}{N} \sum_{i=1}^N \frac{(\hat{y}_i - y_i)}{y_i} \quad [61]$$

$$R^2 = 1 - \frac{\sum_{i=1}^N (\hat{y}_i - y_i)^2}{\sum_{i=1}^N (\hat{y}_i - \bar{y})^2} \quad [62]$$

Onde  $N$  é o número de dados na amostra,  $\hat{y}_i$  é o valor previsto pela rede neural e  $y_i$  é o valor real observado.

#### 4.10. Estatística descritiva

Sumários e apresentações de dados bem constituídos são essenciais ao bom julgamento estatístico, porque permitem ao engenheiro focar nas características importantes dos dados ou ter discernimento acerca do tipo de modelo que deveria ser usado na solução do problema (MONTGOMERY & RUNGER, 2012).

Por esta razão, é importante que se determinem as estatísticas descritivas da amostra de dados experimentais avaliadas em um determinado estudo.

##### 4.10.1 Medidas de tendência central

Representam um ponto em torno do qual os dados se concentram.

A média aritmética é a medida de tendência central mais utilizada para representar os dados. Pode ser entendida como um ponto em torno do qual se concentram os dados.

A média aritmética é dada pela seguinte equação:

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i \quad [63]$$

Onde  $N$  é o número de observações e  $x_i$  é o valor da variável para cada observação.

A assimetria indica que a maior parte dos valores está no centro ou nas extremidades.

A curtose indica o grau de achatamento da distribuição.

##### 4.10.2 Medidas de Dispersão

Segundo Montgomery & Runger (2012), embora a média da amostra seja útil, ela não transmite toda a informação acerca de uma amostra de dados. A variabilidade ou dispersão nos dados pode ser descrita pela variância da amostra ou pelo desvio-padrão da amostra.

Estas medidas representam uma medida do espalhamento dos dados em torno da média da amostra.

A variância de um conjunto de observações é dada pela equação abaixo:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^N (x_i - \bar{x})^2 \quad [64]$$

O Desvio Padrão é dado pela raiz da variância e, deste modo, possui as mesmas unidades da amostra.

$$s = \sqrt{s^2} \quad [65]$$

O coeficiente de variação (CV) é uma medida de dispersão relativa.

$$CV = \frac{s}{\bar{x}} \times 100 \quad [66]$$

## 5. METODOLOGIA

### 5.1. Desenvolvimento do algoritmo de redes neurais

A fim de se realizar o estudo de caso no presente trabalho, foi desenvolvido um algoritmo para o treinamento de uma rede neural em problemas de regressão a partir das equações contidas nos tópicos de propagação da informação e de retropropagação do erro, bem como no tópico do algoritmo de Broyden-Fletcher-Goldfarb-Shanno, com o objetivo de acelerar a convergência do algoritmo para um ponto de mínimo.

O pseudocódigo utilizado segue abaixo:

- 1- Escolher o número de neurônios na camada oculta da rede neural
- 2- Escolher a função de ativação dos neurônios
- 3- Inicializar os pesos sinápticos de forma aleatória
- 4- Inicializar a aproximação da matriz hessiana como a matriz identidade
- 5- Executar o algoritmo de propagação
- 6- Calcular a função custo
- 7- Calcular o gradiente da função custo por meio do algoritmo de retropropagação
- 8- Calcular a direção de descida por meio do algoritmo de BFGS
- 9- Realizar uma pesquisa linear nesta direção a fim de encontrar o passo ótimo
- 10- Encontrado o passo, atualizar os vetores  $s$  e  $y$ , por meio das Equações (55) e (56)
- 11- Atualizar o valor da aproximação da hessiana por meio da Equação (54)
- 12- Repetir a partir do passo 5 até a convergência do algoritmo

O software utilizado no desenvolvimento destes algoritmos foi o R Studio, um software de uso livre.

Segundo seu site oficial, R fornece uma grande variedade de técnicas estatísticas como modelização linear, modelização não linear, testes estatísticos clássicos, análise de séries temporais, classificação e *clustering*, assim como técnicas gráficas, e é bastante extensível.

R está disponível como um software livre sob os termos da *Free Software Foundation's GNU General Public License*. Ele pode ser executado em várias plataformas UNIX e sistemas similares (incluindo FreeBSD e Linux), Windows e MacOS.

O algoritmo desenvolvido para este trabalho é encontrado no anexo A.

## 5.2. Dados experimentais

Um dos objetivos deste trabalho foi o de realizar um estudo de caso acerca do uso de algoritmos de redes neurais na obtenção de um modelo de regressão para a previsão da irradiação solar por meio de variáveis meteorológicas de mais fácil acesso.

Por consequência, dados experimentais contendo as temperaturas máxima e mínima, insolação, velocidade do vento, nebulosidade, pluviosidade e radiação solar foram obtidos junto à Estação Meteorológica da Agronomia da Universidade Federal do Ceará. Foram coletados dados diários correspondentes ao período de 1969 a 2012.

## 5.3. Pré-processamento dos dados experimentais

Além dos dados de temperatura máxima, temperatura mínima, umidade relativa, precipitação total, evaporação, nebulosidade, velocidade do vento e irradiação solar global; dados de irradiação extraterrestre foram utilizados. Estes dados foram calculados para a cidade de Fortaleza a partir da Equação (2).

A partir dos dados experimentais diários, foram gerados três séries de dados: dados diários, dados correspondentes às médias semanais e dados correspondentes às médias mensais de todas as variáveis meteorológicas obtidas.

Em seguida, os dados experimentais foram analisados quanto às suas estatísticas descritivas. Foram, ainda, determinadas o número total de dados disponíveis e a porcentagem de dados em branco a fim de se escolher o melhor conjunto de dados para o estudo de caso.

Finalmente, os dados a serem utilizados foram normalizados de acordo com a equação sugerida por Behrang *et al.* (2010) e encontrada abaixo:

$$X_N = 0,1 + 0,8 * \frac{X_R - X_{Min}}{X_{Máx} - X_{Min}} \quad [66]$$

Onde,  $X_N$  é a variável normalizada,  $X_R$  é o valor real da variável,  $X_{Máx}$  é o valor máximo da variável na amostra de dados e  $X_{Min}$  é o seu valor mínimo.

Segundo LeCun *et al.* (1998), a convergência do algoritmo ocorre de maneira mais rápida se a média de cada variável do conjunto de treinamento for próxima a zero.

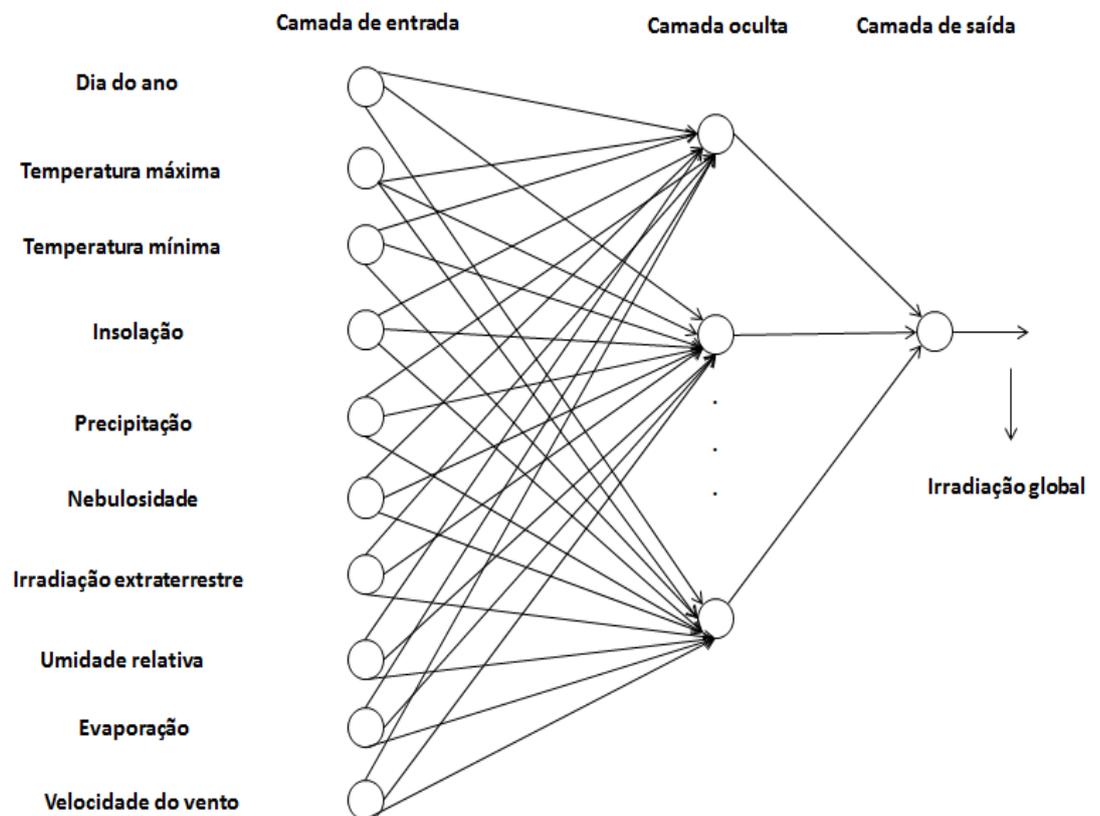
#### 5.4. Estudo de caso

Com as novas variáveis representadas pelos dados meteorológicos normalizados, foram feitos três estudos de caso, correspondentes aos dados diários, semanais e mensais. Os modelos de previsão de irradiação foram desenvolvidos por meio do algoritmo de treinamento de redes neurais desenvolvido para este estudo.

Os modelos estudados neste trabalho consistem de uma rede neural com uma camada oculta tendo como variáveis de entrada as temperaturas máxima e mínima, a umidade relativa, a nebulosidade, a precipitação total, a insolação, a evaporação, a velocidade do vento e a irradiação extraterrestre. Além destas variáveis meteorológicas, utilizou-se o dia do ano, para o grupo de dados diários; o dia da semana, para o grupo de dados semanais e o mês do ano para os dados mensais.

Desta forma, as redes neurais estudadas, possuem a estrutura apresentada na Figura 7.

Figura 7 – Estrutura da rede neural modelada.



Fonte: Elaborada pelo autor.

A fim de testar os diferentes modelos obtidos, os dados foram separados em dois grupos: dados de treinamento e dados de teste. O grupo de dados experimentais corresponde a setenta por cento (70%) da amostra total para cada caso.

Os dados de treinamento foram utilizados para chegar a um modelo de redes neurais e os dados de teste foram utilizados a fim de determinar a capacidade preditiva das redes neurais obtidas quando novos dados que não foram utilizados na fase de treinamento são apresentados a estas redes neurais.

A avaliação da qualidade dos modelos de rede neural foi feita utilizando-se as equações [59] a [62].

A fim de encontrar o número de neurônios nas camadas ocultas que resulte no menor erro de previsão possível para cada estudo de caso, foi desenvolvido um gráfico do comportamento do erro de previsão e do erro de treinamento contra a complexidade das redes neurais, ou seja, contra a quantidade de neurônios ocultos presentes nas redes neurais. Para isto, variou-se o número de neurônios nas camadas ocultas e registraram-se os valores dos erros de previsão e treinamento.

Em seguida, utilizou-se um número fixo de neurônios na camada oculta com o objetivo de determinar o comportamento do erro de previsão com relação à quantidade de dados apresentados às redes. O objetivo desta análise foi obter um conhecimento acerca da quantidade necessária de dados para estudos de previsão de irradiação solar.

Finalmente, utilizou-se o número de neurônios resultante das análises anteriores para treinar as redes neurais novamente e os resultados obtidos avaliados por meio das equações [59] a [61] e apresentados por meio de gráficos e tabelas.

## 6. RESULTADOS E DISCUSSÃO

A variável irradiação solar extraterrestre diária foi calculada a fim de incluí-la como uma das variáveis descritivas do modelo de previsão de irradiação solar global incidente obtida experimentalmente.

Para o conjunto dos dados, as estatísticas descritivas foram determinadas e podem ser encontradas na tabela 2.

Tabela 2 – Porcentagem de dados experimentais presentes em cada ano.

Ano	Tmáx	Tmín	UR1	UR2	UR3	PRECIP	EVAP	INSO	NEB1	NEB2	NEB3	H
1969	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
1970	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
1971	100%	100%	100%	100%	100%	100%	100%	99%	100%	100%	100%	100%
1972	100%	100%	100%	100%	100%	100%	100%	52%	100%	100%	100%	92%
1973	100%	100%	100%	100%	100%	100%	100%	16%	100%	100%	100%	100%
1974	100%	100%	100%	100%	100%	100%	100%	90%	100%	100%	100%	100%
1975	100%	100%	100%	100%	100%	100%	100%	92%	100%	100%	100%	100%
1976	100%	100%	100%	100%	100%	100%	100%	99%	100%	100%	100%	100%
1977	100%	100%	100%	99%	100%	100%	100%	96%	100%	100%	100%	100%
1978	100%	100%	100%	100%	100%	100%	100%	99%	100%	100%	99%	100%
1979	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
1980	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	73%
1981	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
1982	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
1983	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	98%
1984	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
1985	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
1986	100%	100%	100%	100%	100%	100%	100%	99%	100%	100%	100%	100%
1987	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
1988	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%

Fonte: Elaborada pelo autor.

Da Tabela 2, observa-se que o período mais completo é o que vai de 1974 a 1988. Por consequência, este foi o período utilizado em todas as análises deste trabalho.

As estatísticas descritivas para os três grupos de estudo, diário, semanal e mensal encontram-se nas Tabelas 3, 4 e 5.

Tabela 3 – Estatísticas descritivas para o grupo diário.

Variáveis Meteorológicas	H0	TX	TN	INSO	EVAP	UR1	UR2	UR3	PRECIP	NEB1	NEB2	NEB3	H
Média	3599,34	30,19	23,87	7,72	4,08	77,02	70,02	86,75	4,66	6,74	4,75	4,46	1909,44
Desvio-padrão	174,57	1,13	1,21	3,20	1,84	10,54	8,82	6,24	12,42	2,18	2,94	2,98	444,62
CV	5%	4%	5%	42%	45%	14%	13%	7%	267%	32%	62%	67%	23%
Mediana	3626,89	30,30	23,80	9,00	4,00	76,00	69,00	86,00	0,00	7,00	4,00	4,00	2021,88
Mínimo	3184,48	22,40	17,70	0,00	0,00	40,00	37,00	54,00	0,00	0,00	0,00	0,00	115,54
Máximo	3894,42	35,20	29,80	12,00	10,00	100,00	100,00	100,00	198,00	10,00	10,00	10,00	3052,08
Variação	709,94	12,80	12,10	12,00	10,00	60,00	63,00	46,00	198,00	10,00	10,00	10,00	2936,54
Assimetria	-0,46	-0,94	0,02	-1,06	0,36	0,11	0,56	-0,16	5,19	-0,90	0,25	0,14	-1,19
Curtose	-0,57	2,46	-0,03	0,02	0,09	-0,46	0,63	-0,24	38,15	0,65	-1,23	-1,04	1,21

Fonte: Elaborada pelo autor.

Tabela 3 – Estatísticas descritivas para o grupo semanal.

Variáveis Meteorológicas	H0	TX	TN	INSO	EVAP	UR1	UR2	UR3	PRECIP	NEB1	NEB2	NEB3	H
Média	3600,47	30,19	23,87	7,73	4,08	77,02	70,02	86,71	4,63	6,74	4,78	4,47	1909,00
Desvio-padrão	151,52	0,84	1,01	2,15	1,41	8,28	6,70	5,14	6,99	1,35	2,11	1,67	305,76
CV	4%	3%	4%	28%	35%	11%	10%	6%	151%	20%	44%	37%	16%
Mediana	3640,83	30,23	23,86	8,46	4,16	77,21	69,43	86,57	1,49	6,86	4,57	4,29	1950,01
Mínimo	3281,25	27,69	21,14	0,90	0,96	48,07	53,29	73,29	0,00	0,00	0,43	0,43	896,78
Máximo	3803,39	32,34	26,57	10,98	8,59	98,00	87,43	99,29	57,79	9,57	9,86	10,00	2549,36
Variação	522,14	4,65	5,43	10,08	7,63	49,93	34,14	26,00	57,79	9,57	9,43	9,57	1652,58
Assimetria	-0,77	-0,29	0,00	-0,91	-0,02	-0,16	0,18	-0,18	2,44	-0,95	0,24	0,37	-0,68
Curtose	-0,67	-0,34	-0,55	0,06	-0,61	-0,60	-0,51	-0,59	7,95	1,50	-0,89	-0,10	0,16

Fonte: Elaborada pelo autor.

Tabela 3 – Estatísticas descritivas para o grupo mensal.

Variáveis Meteorológicas	H0	TX	TN	INSO	EVAP	UR1	UR2	UR3	PRECIP	NEB1	NEB2	NEB3	H
Média	3601,07	30,18	23,86	7,71	4,08	77,05	70,07	86,76	4,69	6,74	4,77	4,47	1908,91
Desvio-padrão	147,59	0,71	0,91	1,75	1,27	7,34	5,97	4,73	5,17	1,08	1,85	1,17	261,71
CV	4%	2%	4%	23%	31%	10%	9%	5%	110%	16%	39%	26%	14%
Mediana	3641,51	30,20	23,90	8,24	4,08	77,19	70,03	86,80	2,58	6,87	4,67	4,35	1938,64
Mínimo	3305,46	28,33	21,44	2,89	1,28	59,23	56,74	75,10	0,00	2,65	1,67	1,97	1187,45
Máximo	3755,83	31,75	26,14	10,43	7,31	90,13	82,87	96,06	21,79	8,83	9,23	8,17	2461,40
Variação	450,37	3,43	4,70	7,54	6,04	30,90	26,13	20,97	21,79	6,19	7,57	6,20	1273,94
Assimetria	-0,77	-0,21	-0,09	-0,82	-0,04	-0,09	0,09	-0,27	1,32	-0,85	0,22	0,53	-0,49
Curtose	-0,73	-0,61	-0,61	-0,36	-0,66	-0,94	-0,72	-0,57	1,06	0,84	-0,95	0,11	-0,07

Fonte: Elaborada pelo autor.

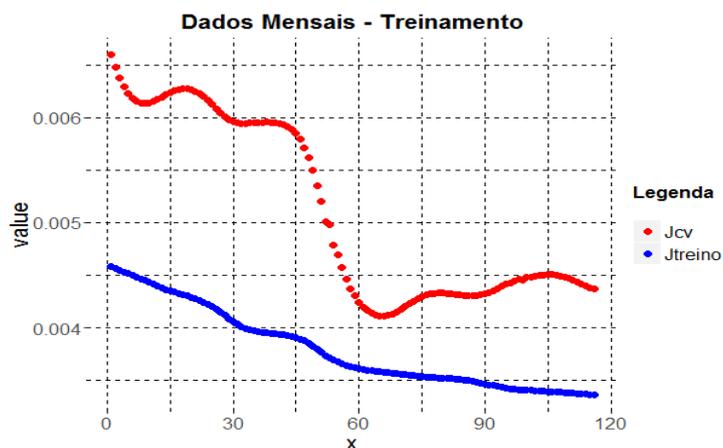
A partir das tabelas acima, percebe-se que não há mudança nos valores das médias das diferentes variáveis meteorológicas quando se varia o período de tempo estudado. Isto ocorre devido à natureza do cálculo da média, que é distribuída entre todos os dados, sendo eles agrupados semanalmente, mensalmente, ou não.

Nota-se, ainda, que o desvio-padrão diminui conforme o período de tempo estudado aumenta. Isto ocorre porque os valores semanais e mensais são obtidos a partir de uma média avaliada em 7 dias e 1 mês, respectivamente. Isto possui o efeito de atenuar a variabilidade da amostra de observações, como representado pela diminuição no coeficiente de variação ao se passar dos dados diários aos dados mensais. Com isto, pode-se criar a hipótese de que um modelo que tenha o objetivo de prever valores de irradiação solar global incidente tenderá a ser menos complexo em termos do número de neurônios para períodos de tempo mais longos. Pode-se supor também que o treinamento de uma rede neural a partir amostras de dados com menor variabilidade necessite de menos amostras de dados para obter um modelo aceitável em termos de erros de previsão e coeficiente de determinação.

Antes de se passar à determinação do número de neurônios com o melhor desempenho para cada grupo de estudo, o resultado de diversos treinamentos foi avaliado graficamente. O objetivo desta análise gráfica foi o de assegurar que o algoritmo desenvolvido para este trabalho está de acordo com a literatura.

O treinamento de uma rede neural visa a minimizar a função custo, ou seja, visa a melhorar o erro de treinamento a cada iteração que ocorra. Dito isto, o comportamento gráfico de alguns treinamentos foi analisado graficamente e o Gráfico 1 foi obtido para o grupo de dados mensais.

Gráfico 1 – Comportamento dos erros de treinamento e teste a cada iteração



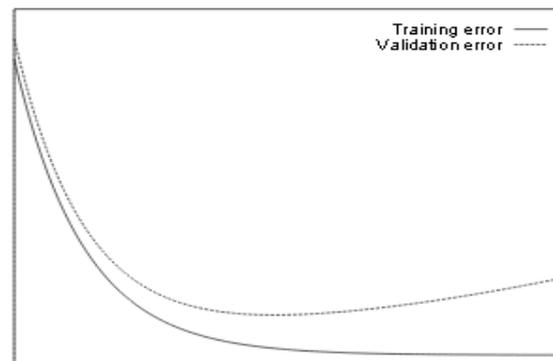
Fonte: Elaborada pelo autor.

Neste gráfico, pode-se observar que a cada avanço no eixo das abscissas, ou seja, a cada iteração, a curva em azul diminui. Isto significa que o erro de treinamento está se dirigindo a um valor menor a cada iteração, exatamente como esperado.

Da mesma forma, é possível observar o comportamento do erro de teste. A curva em vermelho é caracterizada na literatura por começar de modo aleatório (normalmente mais elevada que a de treinamento), diminuir até encontrar um mínimo e em seguida se afastar da curva em azul conforme o erro de treinamento fica cada vez menor. Isto ocorre porque quando a curva de azul está próxima à zero, significa que o modelo criado descreve tão bem os dados de treinamento que ele perde sua capacidade de prever novos pontos.

Segundo Prechelt (1997), o comportamento das curvas de erro de treinamento e teste ideias apresentados na maior parte dos artigos introdutórios sobre redes neurais é representado pela Figura 8.

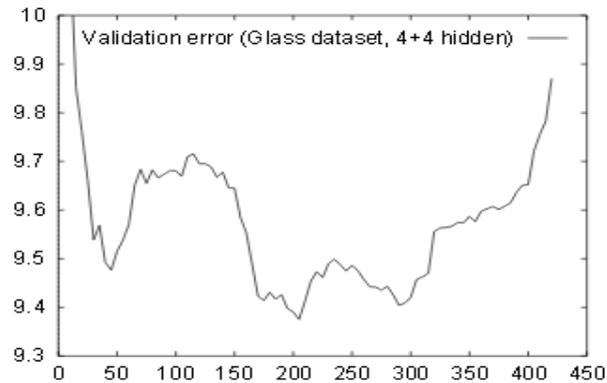
Figura 8 – Erros de treinamento e teste ideias.



Fonte: PRECHELT, 1997.

Entretanto, o erro de teste em aplicações reais de treinamento de uma rede neural não ocorre como uma curva suave como apresentado na Figura 8. Ele ocorre como apresentado na Figura 9. Como pode ser visto, o erro de teste pode apresentar vários picos e vales e em um conjunto de dados realista, é impossível conhecer o erro de teste exato visto que ele só pode ser estimado a partir do conjunto de dados de teste (PRECHELT, 1997).

Figura 9 – Erros de teste reais.



Fonte: PRECHELT, 1997.

Percebe-se, deste modo, que o algoritmo desenvolvido se comporta da maneira esperada e pode-se aplicá-lo ao treinamento das amostras de dados diárias, semanais e mensais.

O próximo passo da análise foi determinar o número da complexidade da rede neural, em termos do número de neurônios da camada oculta, que possibilitaria o menor erro de previsão.

Segundo van der Baan *et al.* (2000), classicamente, a determinação do número de neurônios nas camadas ocultas de uma rede neural se faz por tentativa e erro. Diferentes estruturas são treinadas e têm seus desempenhos avaliados. Finalmente, a melhor configuração é escolhida. O problema deste tipo de procedimento é seu alto custo manual, embora algoritmos computacionais possam ser escritos a fim de construir, treinar e testar o desempenho de diversas redes neurais.

Os resultados obtidos para os diferentes grupos de estudo são mostrados nos gráficos 2, 3 e 4 a seguir:

Gráfico 2 – Erro versus Complexidade da rede para o grupo diário

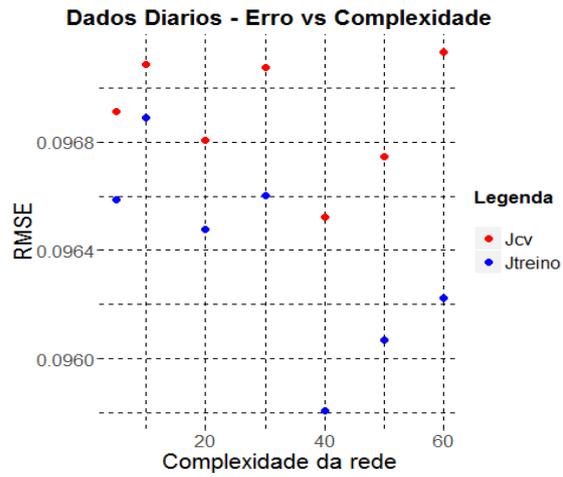


Gráfico 3 – Erro versus Complexidade da rede para o grupo semanal

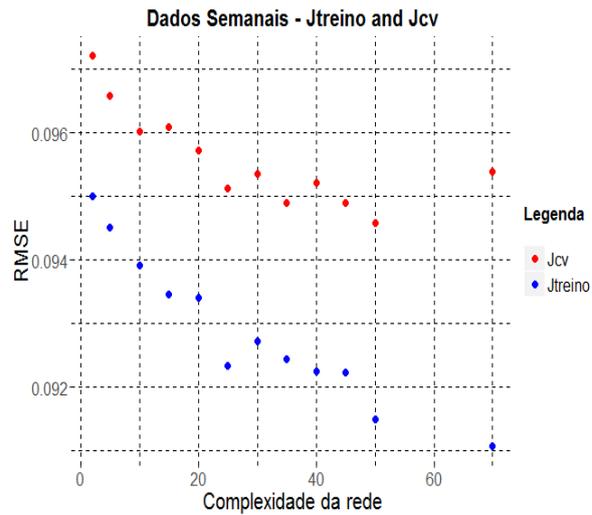
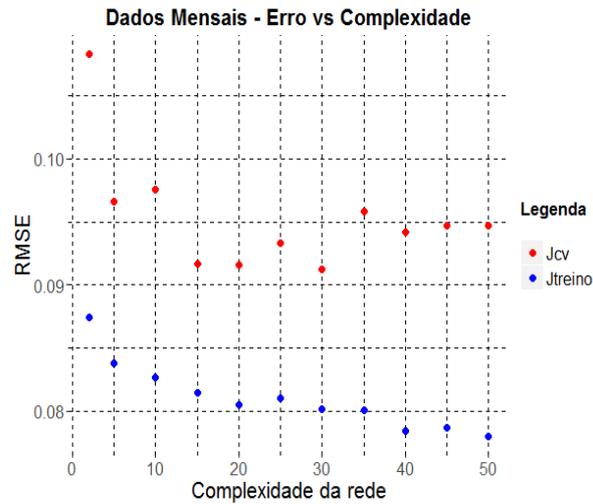


Gráfico 4 – Erro versus Complexidade da rede para o grupo mensal

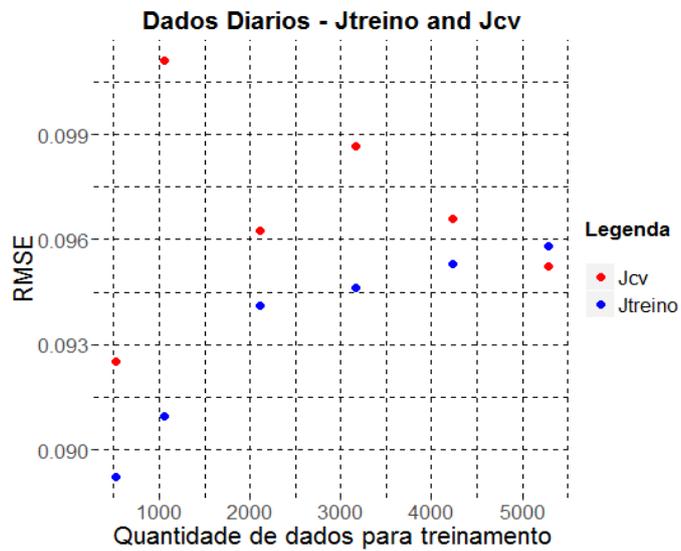


Fonte: Elaborada pelo autor.

O Gráfico 4 mostra que para o grupo de dados mensais, o número de neurônios na camada oculta que torna o erro de previsão mínimo é de 30 neurônios. A partir deste número de neurônios, a rede descreve muito bem os dados de treinamento, mas perde sua capacidade de generalizar o modelo para descrever dados novos. Antes de 30 neurônios, a rede não descreve bem os dados de treinamento e não generaliza o modelo de modo satisfatório, ou seja, a rede neural não aprendeu a descrever a irradiação solar incidente mensal. De maneira análoga, o número de neurônios de melhor capacidade preditiva é de 40 e 50, para os dados diários e semanais, respectivamente.

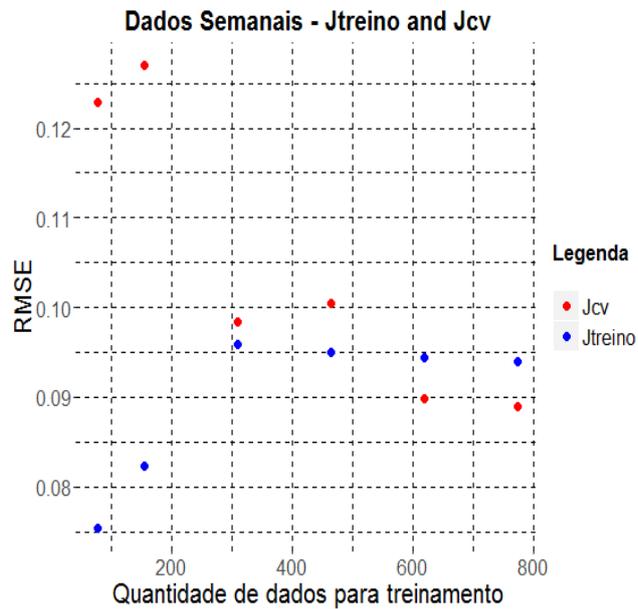
A seguir, analisou-se o efeito da quantidade de dados utilizada no treinamento da rede neural. Os resultados estão sintetizados nos gráficos 5, 6 e 7 a seguir:

Gráfico 5 – Erro de treinamento e teste versus Quantidade de dados para treinamento para o grupo diário



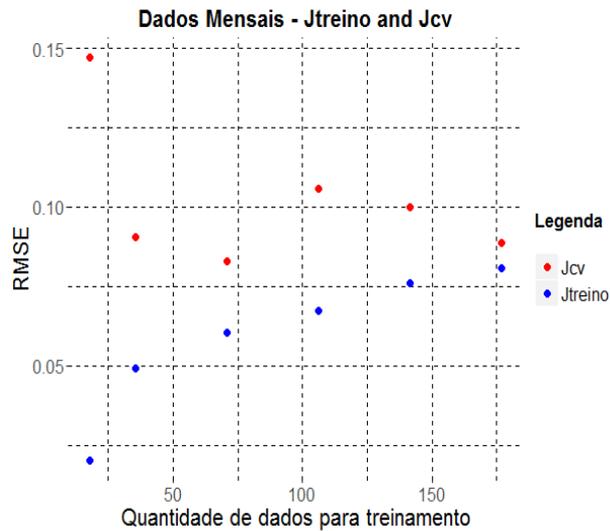
Fonte: Elaborada pelo autor.

Gráfico 6 – Erro de treinamento e teste versus Quantidade de dados para treinamento para o grupo semanal.



Fonte: Elaborada pelo autor.

Gráfico 7 – Erro de treinamento e teste versus Quantidade de dados para treinamento para o grupo mensal.

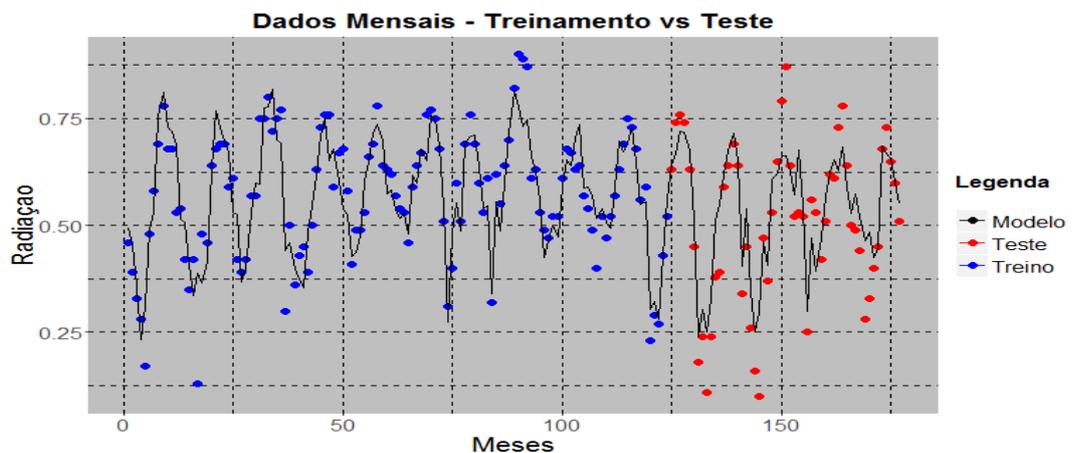


Fonte: Elaborada pelo autor.

A análise destes gráficos torna clara a importância de se obter uma amostra considerável de dados a serem utilizados durante o treinamento e o teste das redes neurais. É possível confirmar que, de um modo geral, o aumento do número de dados utilizados, as redes neurais conseguem generalizar melhor, diminuindo o erro de previsão. Em outras palavras, as redes neurais aprendem cada vez mais e melhor a medida que encontram mais dados.

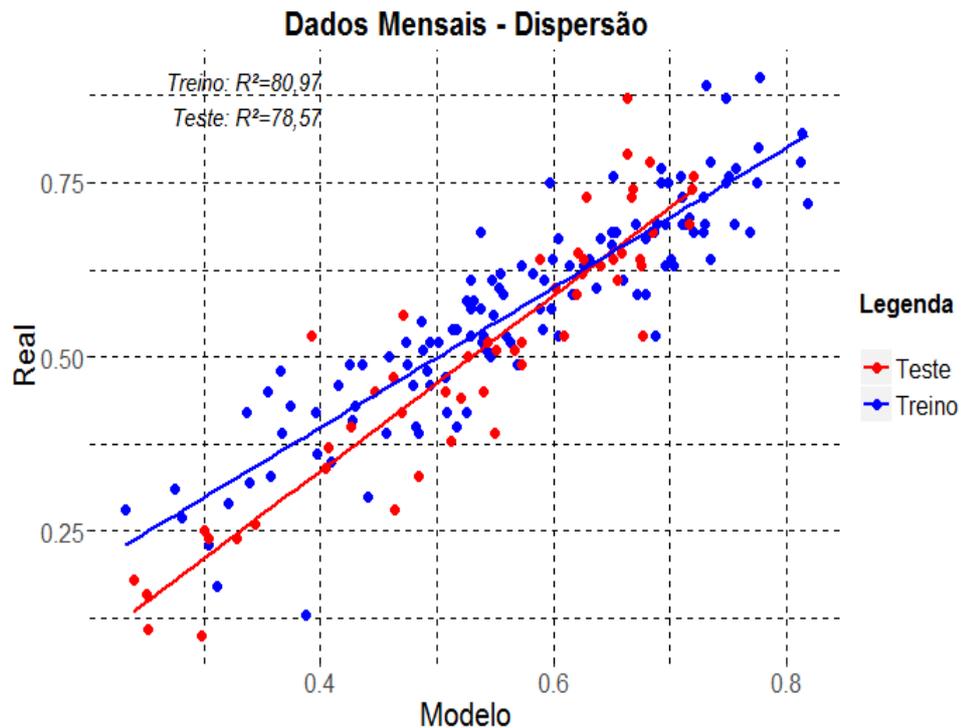
Por fim, com base no número de neurônios obtidos anteriormente, uma rede neural foi criada para cada estudo de caso e os resultados podem ser observados nos gráficos 8 e 9 e na Tabela 5 a seguir:

Gráfico 8 – Dados de treinamento, dados de teste de modelo obtido para o grupo mensal.



Fonte: Elaborada pelo autor.

Gráfico 9 – Dispersão de treinamento e teste para o grupo mensal



Fonte: Elaborada pelo autor.

Tabela 5 – Coeficientes de determinação dos modelos

	Treino				Teste			
	RMSE	RRMSE (%)	MAPE (%)	$R^2$	RMSE	RRMSE(%)	MAPE (%)	$R^2$
Diários	0,952	15,57	14,59	35,67	8,76	6,29	14,87	50,26
Semanais	0,0828	13,68	12,51	64,98	0,0929	6,57	18,17	65,73
Mensais	0,0655	11,47	11,09	80,97	0,0858	6,42	20,83	78,58

Fonte: Elaborada pelo autor.

Segundo YADAV *et al.* (2014), um MAPE menor que 10%, indica uma excelente capacidade preditiva; um MAPE entre 10% e 20%, representa um modelo com boa capacidade de previsão; um MAPE entre 20% e 50%, indica uma capacidade preditiva razoável e um modelo com MAPE acima de 50% representa um modelo ineficaz.

A partir deste critério e dos resultados finais obtidos, pode-se concluir que os modelos gerados com o algoritmo de redes neurais são capazes de obter bons resultados.

Segundo Li *et al.* (2013), diferentes intervalos de RRMSE podem ser definidos a fim de demonstrar a capacidade de modelos como: excelente (RRMSE menor que 10%), bom (RRMSE entre 10% e 20%), razoável (RRMSE entre 20% e 30%) e ineficaz (RRMSE maior que 30%).

Desta forma, os modelos desenvolvidos neste trabalho se encontram na classificação excelente quanto às suas capacidades preditivas tendo em vista esse critério.

Shamshirband *et al.* (2015), estudaram diferentes modelos baseados em inteligência artificial para a previsão da média mensal de irradiação solar global a partir de variáveis meteorológicas para a cidade iraniana de Abbas. Seus resultados mostram que no conjunto de dados de testes, o coeficiente de determinação variou de 54,87 a 97,37.

Dito isto, o modelo para a previsão da média mensal de irradiação global incidente desenvolvido neste trabalho possui um valor do coeficiente de determinação dentro deste intervalo.

Os modelos criados são mais consistentes para os casos semanais e mensais. Uma explicação para este fenômeno pode ser o aumento da variabilidade do conjunto de dados experimentais ao se passar de um modelo mensal para um semanal e, por fim, diário. Este aumento na variabilidade dos dados, percebido pelo aumento do coeficiente de variação (CV), pode dificultar a capacidade de um modelo descrever bem um conjunto de dados mais dispersos, diminuindo seu coeficiente de determinação.

## 7. CONCLUSÃO

O presente trabalho consistiu em desenvolver um algoritmo de treinamento de redes neurais e aplicá-lo no estudo da previsão de irradiação global incidente a partir de variáveis meteorológicas. Para isto, um algoritmo de treinamento de redes neurais foi desenvolvido e otimizado por meio do algoritmo de BFGS utilizando-se o software R Studio. Além disto, foi utilizado um histórico de dados experimentais contendo valores das variáveis meteorológicas para o período de 1974 a 1988. Portanto, o desenvolvimento do algoritmo de redes neurais, sua otimização por meio do algoritmo BFGS, assim como a coleta de dados experimentais foram realizados com sucesso.

Os dados experimentais foram analisados quanto às suas estatísticas descritivas e foram normalizados a fim de prepara-los para o uso no estudo de caso. Desta forma, o pré-processamento dos dados experimentais foi realizado com sucesso.

Por fim, foi possível criar redes neurais para a previsão de irradiação solar global em três estudos de caso: previsão da irradiação solar global diária, média semanal da irradiação global e média mensal da irradiação global. Da análise, dos coeficientes de determinação para os diferentes estudos de caso, concluiu-se que o modelo para previsão da média mensal de irradiação global foi o melhor. Portanto, a aplicação dos algoritmos desenvolvidos no estudo da previsão de irradiação solar foi realizada com sucesso.

## 8. REFERÊNCIAS

- BEHRANG, M. A.; ASSAREH, E.; GHANBARZADEH, A.; NOGHREHABADI, A. R. The potential of different artificial neural network (ANN) techniques in daily global solar radiation modeling based on meteorological data. **Solar Energy** 2010. 84: 1468-1480.
- BENGHANEM, M.; MELLIT, A.; ALAMRI, S.N. ANN-based modeling and estimation of daily global solar radiation data: a case study. **Energy Conversion and Management** 2009. 50: 1644-1655.
- BISHOP, Christopher M. Pattern recognition and machine learning. 1ª Edição. New York, Springer: 2007.
- BOSCH, J.L., LOPEZ, G., BATLLES, F.J. Daily solar irradiation estimation over a mountainous area using artificial neural networks. **Renewable Energy** 2008. 33: 1622-1628.
- COPPIN, Ben. Inteligência artificial. 1ª Edição. Rio de Janeiro, LTC: 2010.
- DAI, Yu-Hong. Convergence properties of the BFGS algorithm. **Society for Industrial and Applied Mathematics** 2002. 13: 693-701.
- DORESWAMY and CHANABASAYYA. M. Vastrad. Performance analysis of neural network models for oxazolines and oxazoles derivatives descriptor dataset. **International Journal of Information Sciences and Techniques (IJIST)** 2013. Vol.3, No. 6.
- DUFFIE, John A.; BECKMAN, William A. Solar engineering of thermal processes. 4ª Edição. New Jersey. Wiley: 2013.
- FACELI, Katti; LORENA, Ana C.; GAMA, João; DE CARVALHO, André C. P. L. F. Inteligência artificial: uma abordagem de aprendizado de máquina. 1ª Edição. Rio de Janeiro, LTC: 2011.
- HAYKIN, Simon. Redes neurais: princípios e prática. 2ª Edição. Porto Alegre. Bookman: 2001.
- HUNT, L.; KUCHAR, L.; SWANTON, C. Estimation of solar radiation for use in crop modeling. **Agricultural and Forest Meteorology** 1998. 91: 293-300
- MONTGOMERY, Douglas C.; RUNGER, George C. Estatística aplicada e probabilidade para engenheiros. 4ª Edição. Rio de Janeiro. LTC: 2009.
- KHATIB, Tamer, MOHAMED, Azah, SOPIAN, K. A review of solar energy modeling techniques. **Renewable and Sustainable Energy Reviews** 2012, 16: 2864 – 2869.
- KISI, Ozgur. Modeling solar radiation of Mediterranean region in Turkey by using fuzzy genetic approach. **Energy** 2014; 64: 429–436.
- LECUN, Yann, BOTTOU, Leon, ORR, Genevieve B.; MULLER, Klaus-Robert. Efficient backprop. In: MONTAVON, Grégoire; ORR, Geneviève; MULLER, Klaus-Robert. **Neural networks: tricks of the trade**. New York: Springer, 2012. P. 9-48.

LI, Mao-Fen, TANG, XIAO-Ping, WU, Wei, LIU, Hong-Bin. General models for estimating daily global solar radiation for different solar radiation zones in mainland China. **Energy Conversion Management** 2013. 70: 139-148.

MARQUES FILHO, Edson P.; OLIVEIRA, Amauri P.; VITA, Willian A.; MESQUITA, Francisco L.L.; CODATO, Georgia, ESCOBEDO, João F.; CASSOL, Mariana, FRANÇA, José Ricardo A. Global, diffuse and direct solar radiation at the surfasse in the city of Rio de Janeiro: Observational characterization and empirical modeling. **Renewable Energy** 2016. 91: 64-74.

MOHAMMADI, Kasra, SHAMSHIRBAND, Shahaboddin, TONG, Chong W., ALAM, Khubaib A.; PETKOVIC, Dalibor. Potential of adaptive neuro-fuzzy system for prediction of daily global solar radiation by day of the year. **Energy Conversion and Management** 2015, 93: 406 – 413.

MOHAMMADI, Kasra, SHAMSHIRBAND, Shahaboddin, TONG, Chong W.; ARIF, Muhammad, PETKOVIC, Dalibor, CH, Sudheer. A new hybrid support vector machine-wavelet transform approach for estimation of horizontal global solar radiation. **Energy Conversion and Management** 2015, 92: 162 – 171.

OLATOMIWA, Lanre, MEKHILEF, Saad, SHAMSHIRBAND, Shahaboddin, MOHAMMADI, Kasra, PETKOVIC, Dalibor, CH, Sudheer. A support vector machine-firefly algorithm-based model for global solar radiation prediction. **Solar Energy** 2015. 115: 632-644.

PENG, Chun-Cheng, MAGOULAS, George D. Nomonotone BFGS-trained recurrent neural networks for temporal sequence processing. **Applied Mathematics and Computation** 2011. 217: 5421-5441.

PRECHELT, Lutz. Early stop – but when?. In: MONTAVON, Grégoire; ORR, Geneviève; MULLER, Klaus-Robert. **Neural networks: tricks of the trade**. New York: Springer, 2012. P. 53-67.

SHAMSHIRBAND, Shahaboddin, MOHAMMADI, Kasra, TONG, Chong W.; ZAMANI, Mazdak, MOTAMEDDI, Shervin, CH, Sudheer. A hybrid SVM-FFA method for prediction of monthly mean global solar radiation. **Spring-Verlag Wien** 2015.

VAN DER BAAN, M., JUTTEN, C. Neural networks in geophysical applications. **Geophysics** 2000, 65: 1032–1047.

YADAV, Amit K.; MALIK, Hasmat, S.S., Chandel. Selection of most relevant input parameters using WEKA for artificial neural network based solar radiation prediction models. **Renewable and Sustainable Energy Reviews** 2014, 31: 509 – 519.