



**Universidade do Estado do Rio de Janeiro**

Centro de Tecnologia e Ciências

Instituto de Matemática e Estatística

Oswaldo Areal


**Inferência de frações de butano no propano e de frações de  
propano no butano em uma coluna depropanizadora utilizando  
rede neuronal**

Rio de Janeiro

2017

Oswaldo Areal

**Inferência de frações de butano no propano e de frações de propano no butano em uma coluna depropanizadora utilizando rede neuronal**



Dissertação apresentada, como requisito parcial para obtenção do título de Mestre, ao Programa de Pós-Graduação em Ciências Computacionais, da Universidade do Estado do Rio de Janeiro.

Orientadora: Prof.<sup>a</sup> Dra. Roseli Suzi Wedemann  
Coorientadora: Prof.<sup>a</sup> Dra. Patrícia Nunes da Silva

Rio de Janeiro

2017

CATALOGAÇÃO NA FONTE  
UERJ / REDE SIRIRUS / BIBLIOTECA CTC-A

A678

Areal, Oswaldo.

Inferência de frações de butano no propano e de frações de propano no butano em uma coluna depropanizadora utilizando Rede Neuronal / Oswaldo Areal. – 2017.

78f.: il.

Orientadora: Roseli Suzi Wedemann.

Coorientadora: Patrícia Nunes da Silva.

Dissertação (Mestrado Ciências Computacionais) - Universidade do Estado do Rio de Janeiro, Instituto de Matemática e Estatística.

1. Redes neurais (Computação) - Teses. 2. Petróleo - Derivados - Teses. 3. Propano - Teses. 4. Butano - Teses. I. Wedemann, Roseli Suzi. II. Silva, Patrícia Nunes da. III. Universidade do Estado do Rio de Janeiro. Instituto de Matemática e Estatística. IV. Título.

CDU 007.52

Autorizo para fins acadêmicos e científicos, a reprodução total ou parcial desta dissertação, desde que citada a fonte.

---

Assinatura

---

Data

Oswaldo Areal

**Inferência de frações de butano no propano e de frações de propano no butano em uma coluna depropanizadora utilizando rede neuronal**

Dissertação apresentada, como requisito parcial para obtenção do título de Mestre, ao Programa de Pós-Graduação em Ciências Computacionais, da Universidade do Estado do Rio de Janeiro.

Aprovada em 11 de Março de 2017.

Banca Examinadora:

---

Prof.<sup>a</sup> Dra. Roseli Suzi Wedemann (Orientadora)  
Instituto de Matemática e Estatística – UERJ

---

Prof.<sup>a</sup> Dra. Patrícia Nunes da Silva (Coorientadora)  
Instituto de Matemática e Estatística – UERJ

---

Prof. Dr. Carlos Frederico Fragoso de Barros e Vasconcellos  
Universidade do Estado do Rio de Janeiro

---

Prof. Dr. Luís Alfredo Vidal de Carvalho  
Universidade Federal do Rio de Janeiro

Rio de Janeiro

2017

## DEDICATÓRIA

Dedico este trabalho aos meus pais, Gentil e Oswaldina, que com esforço e abnegação fizeram que eu tenha formação acadêmica, e a minha família, Solange e Nathália que muito contribuíram com compreensão e carinho.

## AGRADECIMENTOS

A Deus que me permitiu chegar até aqui.

As minhas orientadoras, Roseli e Patrícia, pelo tempo e dedicação para a realização deste trabalho.

Ao meu colega de trabalho Manuel Brana que me incentivou a cursar o mestrado.

Aos professores e funcionários do IME que contribuíram direta ou indiretamente na minha formação.

Por vezes sentimos que aquilo que fazemos não é senão uma gota de água no mar. Mas o mar seria menor se lhe faltasse uma gota.

*Madre Teresa de Calcutá*

## RESUMO

AREAL, O. F. *Inferência de frações de butano no propano e de frações de propano no butano em uma coluna depropanizadora utilizando rede neuronal*. 2017. 78 f. Dissertação (Mestrado em Ciências Computacionais) – Instituto de Matemática e Estatística, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2017.

Parâmetros de qualidade (especificações) de derivados de petróleo são dependentes de variáveis de processo, tais como vazões, pressões e temperaturas. Entretanto não se conhece a expressão analítica da função que expressa essa dependência. Mas, essa dependência está registrada nos dados coletados ao longo dos anos, durante o funcionamento da cadeia produtiva de uma refinaria. Este trabalho apresenta os estudos realizados para a construção de uma rede neuronal, com o objetivo de calcular parâmetros de qualidade de derivados de petróleo. A base de dados permite que as redes neuronais sejam utilizadas para estimar, em tempo real, os parâmetros de qualidade de interesse, em função das variáveis de processo. A arquitetura da rede neuronal utilizada foi a perceptron de múltiplas camadas. As especificações escolhidas foram os teores de frações pesadas na produção de topo e os de frações leves na produção de fundo de uma coluna de destilação. Dentro da arquitetura escolhida foram testadas varias topologias e comparados os desempenhos em função do erro quadrático médio, com o objetivo de encontrar a topologia mais adequada. Apesar de ter sido provado que uma rede neuronal é capaz de aproximar qualquer função contínua, encontrar a arquitetura e a topologia adequadas para a rede constitui-se um grande desafio.

Palavras-chave: Rede neuronal. Coluna depropanizadora. Propano. Butano.



## ABSTRACT

AREAL, O. F. *Inference of butane fractions in propane and propane fractions in butane, in a depropanizer column by neural network.* 2017. 78 f. Dissertação (Mestrado em Ciências Computacionais) – Instituto de Matemática e Estatística, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2017.

This paper presents the studies carried out for the construction of a neural network, in order to calculate quality parameters (specifications) of oil derivatives. These quality parameters are dependent on process parameters, such as flows, pressures and temperatures, but the analytical expression of the function expressing this dependence is not known. However, this dependence is registered in the data collected throughout the years, during the operation of the production chain the refinery. This database allows the neural networks to be used to estimate, in real time, the quality parameters of interest as a function of the process parameters. The architecture of the neural network used was the multilayer perceptron. The chosen specifications were the percentage of heavy fractions in the top production and the light fractions in the bottom production of a distillation column. With the chosen architecture, several topologies were tested and the performance in function of the mean square error, in order to find the most appropriate topology. Although it has been proven that a neural network is able to approximate any continuous function, finding the appropriate architecture and topology for the network is a great challenge.

Keywords: Neural network. Depropanizer column. Propane. Butane.

## LISTA DE ILUSTRAÇÕES

Figura 1 - Esquema de Produção de Derivados de uma Refinaria. . . . .	17
Figura 2 - Esquema de uma Unidade de Craqueamento Catalítico. . . . .	18
Figura 3 - Coluna Depropanizadora. . . . .	19
Figura 4 - Prato com borbulhadores. . . . .	19
Figura 5 - Mudanças de potencial do neurônio durante sua excitação. . . . .	21
Figura 6 - Função de Ativação. . . . .	23
Figura 7 - Função degrau. . . . .	24
Figura 8 - Modelo de um neurônio artificial. . . . .	25
Figura 9 - Perceptron simples. . . . .	26
Figura 10 - Perceptron de múltiplas camadas. . . . .	26
Figura 11 - Rede recorrente. . . . .	27
Figura 12 - Cosseno e a saída produzida pela rede neuronal sem convergir com topologia de uma unidade na camada de entrada, 30 neurônios na primeira camada oculta, 20 neurônios na segunda camada, 10 na terceira, 5 na quarta e um neurônio na camada de saída. . . . .	32
Figura 13 - Cosseno e a saída da rede com topologia de uma unidade na camada de entrada, 2 neurônios na camada oculta e um neurônio na camada de saída. . . . .	34
Figura 14 - Cosseno e a saída da rede com topologia de uma unidade na camada de entrada, 3 neurônios na camada oculta e um neurônio na camada de saída. . . . .	35
Figura 15 - Cosseno e a saída da rede com topologia de uma unidade na camada de entrada, 4 neurônios na camada oculta e um neurônio na camada de saída. . . . .	36
Figura 16 - Cosseno e a saída da rede com topologia de uma unidade na camada de entrada, 5 neurônios na camada oculta e um neurônio na camada de saída. . . . .	37
Figura 17 - Evolução do erro em função do aumento de neurônios. . . . .	38
Figura 18 - Cosseno e a saída da rede com topologia de uma unidade na camada de entrada, 5 neurônios na primeira camada oculta, 2 neurônios na segunda camada e um neurônio na camada de saída. . . . .	39
Figura 19 - Cosseno e a saída da rede com topologia de uma unidade na camada de entrada, 5 neurônios na primeira camada oculta, 3 neurônios na segunda camada e um neurônio na camada de saída . . . . .	40

Figura 20 - Cosseno e a saída da rede com topologia de uma unidade na camada de entrada, 5 neurônios na primeira camada oculta, 4 neurônios na segunda camada e um neurônio na camada de saída . . . . .	41
Figura 21 - Seno e a saída da rede com topologia de uma unidade na camada de entrada, 7 neurônios na primeira camada oculta, 5 neurônios na segunda camada e um neurônio na camada de saída . . . . .	42
Figura 22 - Erro da rede em função da topologia com duas camadas ocultas e intervalo na entrada de $[-9.4, 10.7]$ . . . . .	43
Figura 23 - Cosseno e a saída da rede com topologia de uma unidade na camada de entrada, 14 neurônios na primeira camada oculta, 10 neurônios na segunda camada e um neurônio na camada de saída . . . . .	45
Figura 24 - Erro da rede em função da topologia com duas camadas ocultas e intervalo na entrada de $[-15.708, 31.41]$ . . . . .	45
Figura 25 - Erro da rede em função da topologia com três camadas ocultas e intervalo na entrada de $[-15.708, 31.41]$ . . . . .	47
Figura 26 - Cosseno e a saída da rede com topologia de uma unidade na camada de entrada, 15 neurônios na primeira camada oculta, 10 neurônios na segunda, 5 na terceira e um neurônio na camada de saída . . . . .	47
Figura 27 - Cosseno e a saída da rede com topologia de uma unidade na camada de entrada, 20 neurônios na primeira camada oculta, 15 neurônios na segunda camada, 10 na terceira, 5 na quarta e um neurônio na camada de saída . . . . .	48
Figura 28 - Exemplo de topologia dos experimentos . . . . .	53
Figura 29 - Saída da rede neuronal e o valor esperado de percentual de frações pesadas no propano, em relação à razão de refluxo de topo e produção de propano (R.R.P em %), com topologia de 10 unidades na camada de entrada, 10 neurônios na camada oculta e 2 neurônios na camada de saída . . . . .	54
Figura 30 - Saída da rede neuronal e o valor esperado de percentual de frações leves no butano, em relação à razão de refluxo de topo e produção de propano (R.R.P em %), com topologia de 10 unidades na camada de entrada, 10 neurônios na camada oculta e 2 neurônios na camada de saída . . . . .	55
Figura 31 - Saída da rede neuronal e o valor esperado de percentual de frações pesadas no propano, em relação à razão de refluxo de topo e produção de propano (R.R.P em %), com topologia de 10 unidades na camada de entrada, 15 neurônios na camada oculta e 2 neurônios na camada de saída. . . . .	56

Figura 32 - Saída da rede neuronal e o valor esperado de percentual de frações leves no butano, em relação à razão de refluxo de topo e produção de propano (R.R.P em %), com topologia de 10 unidades na camada de entrada, 15 neurônios na camada oculta e 2 neurônios na camada de saída. . . . .	56
Figura 33 - Saída da rede neuronal e o valor esperado de percentual de frações pesadas no propano, em relação à razão de refluxo de topo e produção de propano (R.R.P em %), com topologia de 10 unidades na camada de entrada, 20 neurônios na camada oculta e 2 neurônios na camada de saída. . . . .	57
Figura 34 - Saída da rede neuronal e o valor esperado de percentual de frações leves no butano, em relação à razão de refluxo de topo e produção de propano (R.R.P em %), com topologia de 10 unidades na camada de entrada, 20 neurônios na camada oculta e 2 neurônios na camada de saída. . . . .	58
Figura 35 - Saída da rede neuronal e o valor esperado de percentual de frações pesadas no propano, em relação à razão de refluxo de topo e produção de propano (R.R.P em %), com topologia de 10 unidades na camada de entrada, 15 neurônios na camada oculta e um neurônio na camada de saída. . . . .	59
Figura 36 - Saída da rede neuronal e o valor esperado de percentual de frações pesadas no propano, em relação à razão de refluxo de topo e produção de propano (R.R.P em %), com topologia de 10 unidades na camada de entrada, 30 neurônios na primeira camada oculta, 15 neurônios na segunda camada e um neurônio na camada de saída. . . . .	61
Figura 37 - Saída da rede neuronal e o valor esperado de percentual de frações pesadas no propano, em relação à razão de refluxo de topo e produção de propano (R.R.P em %), com topologia de 10 unidades na camada de entrada, 30 neurônios na primeira camada oculta, 15 neurônios na segunda camada, 5 na terceira e um neurônio na camada de saída. . . .	62
Figura 38 - Saída da rede neuronal e o valor esperado de percentual de frações pesadas no propano, em relação à razão de refluxo de topo e produção de propano (R.R.P em %), com topologia de 9 unidades na camada de entrada, 30 neurônios na primeira camada oculta, 15 neurônios na segunda camada, 5 na terceira e um neurônio na camada de saída. . . .	63

Figura 39 - Saída da rede neuronal e o valor esperado de percentual de frações pesadas no propano, em relação à razão de refluxo de topo e produção de propano (R.R.P em %), com topologia de 8 unidades na camada de entrada, 30 neurônios na primeira camada oculta, 15 neurônios na segunda, 5 na terceira e um neurônio na camada de saída. . . . .	64
Figura 40 - Saída da rede neuronal e o valor esperado de percentual de frações pesadas no propano, em relação à razão de refluxo de topo e produção de propano (R.R.P em %), com topologia de 6 unidades na camada de entrada, 30 neurônios na primeira camada oculta, 15 neurônios na segunda, 5 na terceira e um neurônio na camada de saída. . . . .	65
Figura 41 - Saída da rede neuronal e o valor esperado de percentual de frações pesadas no propano, em relação à razão de refluxo de topo e produção de propano (R.R.P em %), com topologia de 5 unidades na camada de entrada, 30 neurônios na primeira camada oculta, 15 neurônios na segunda camada, 5 na terceira e um neurônio na camada de saída. . .	66
Figura 42 - Saída da rede neuronal e o valor esperado de percentual de frações pesadas no propano, em relação à razão de refluxo de topo e produção de propano (R.R.P em %). . . . .	67
Figura 43 - Saída da rede neuronal e o valor esperado de percentual de frações pesadas no propano, em relação à razão de refluxo de topo e produção de propano (R.R.P em %), com topologia de 6 unidades na camada de entrada, 30 neurônios na primeira camada oculta, 15 neurônios na segunda camada, 5 na terceira e um neurônio na camada de saída. . .	68

## LISTA DE TABELAS

Tabela 1 - Pesos finais. . . . .	38
Tabela 2 - Topologia de duas camadas e intervalo da entrada de $[-9.4, 10.7]$ . . .	42
Tabela 3 - Topologia de duas camadas e intervalo de entrada de $[-15.708, 31.41]$ . . .	44
Tabela 4 - Topologia de três camadas. . . . .	46
Tabela 5 - Variáveis de entrada. . . . .	50
Tabela 6 - Variáveis de entrada após tratamento. . . . .	52
Tabela 7 - Variáveis de entrada do melhor experimento. . . . .	69

## SUMÁRIO

	<b>INTRODUÇÃO</b>	15
1	<b>DERIVADOS DE PETRÓLEO</b>	17
1.1	Coluna Depropanizadora	18
2	<b>REDE NEURONAL ARTIFICIAL</b>	21
2.1	Neurônio Biológico	21
2.2	Neurônio Artificial	22
2.3	Função de Ativação	22
2.3.1	Saída do Neurônio	24
2.4	Redes Neurais Artificiais	25
2.5	Retropropagação do Erro	28
2.6	Algoritmo de Retropropagação de Erros - Backpropagation	28
3	<b>EXPERIMENTOS PRELIMINARES</b>	30
3.1	Topologia	30
3.2	Derivada da Função de Ativação	31
3.3	Teste inicial	31
3.4	Resultados	33
3.4.1	<u>Uma camada oculta.</u>	33
3.4.1.1	Uma camada oculta com dois neurônios.	33
3.4.1.2	Uma camada oculta com três neurônios.	34
3.4.1.3	Uma camada oculta com quatro neurônios.	35
3.4.1.4	Uma camada oculta com cinco neurônios.	36
3.4.2	<u>Duas camadas ocultas.</u>	39
3.4.2.1	Duas camadas ocultas com cinco e dois neurônios.	39
3.4.2.2	Duas camadas ocultas com cinco e três neurônios.	39
3.4.2.3	Duas camadas ocultas com cinco e quatro neurônios.	40
3.4.2.4	Duas camadas ocultas com sete e cinco neurônios.	41
3.4.3	<u>Três camadas ocultas.</u>	46
3.4.3.1	Três camadas ocultas com quinze, dez e cinco neurônios.	46
3.4.4	<u>Quatro camadas ocultas.</u>	48
3.4.4.1	Quatro camadas ocultas com vinte, quinze, dez e cinco neurônios.	48
4	<b>A COLUNA DEPROPANIZADORA</b>	50
4.1	Tratamento inicial dos dados.	51
4.2	Resultados	52
4.3	<b>Uma camada oculta e dois neurônios na camada de saída.</b>	53
4.3.1	<u>Uma camada oculta com dez neurônios.</u>	54
4.3.2	<u>Uma camada oculta com quinze neurônios.</u>	55

4.3.3	<u>Uma camada oculta com vinte neurônios.</u>	56
4.4	<b>Uma camada oculta e um neurônio na saída.</b>	59
4.4.1	<u>Uma camada oculta com quinze neurônios.</u>	59
4.4.2	<u>Uma camada oculta com trinta neurônios.</u>	60
4.5	<b>Dois camadas ocultas e um neurônio na saída.</b>	60
4.5.1	<u>Dois camadas ocultas com trinta e quinze neurônios.</u>	60
4.6	<b>Três camadas ocultas e um neurônio na saída.</b>	61
4.6.1	<u>Três camadas ocultas com trinta, quinze e cinco neurônios.</u>	61
4.7	<b>Retirando variáveis de entrada da rede.</b>	62
4.7.1	<u>Rede com nove unidades de entrada.</u>	63
4.7.2	<u>Rede com oito unidades de entrada.</u>	64
4.7.3	<u>Rede com seis unidades de entrada.</u>	65
4.7.4	<u>Rede com cinco unidades de entrada.</u>	66
4.7.5	<u>Rede após troca de variável de entrada.</u>	67
4.7.6	<u>Rede com seis unidades de entrada, após inclusão de variáveis</u>	68
4.8	<b>Experimentos com dados próximos da média</b>	69
4.8.1	<u>Novo tratamento nos dados de entrada</u>	69
4.8.2	<u>Variáveis de entrada do melhor experimento</u>	69
	<b>CONCLUSÕES</b>	71
	<b>REFERÊNCIAS</b>	75
	<b>ANEXO A – Pseudocódigo</b>	76



## INTRODUÇÃO

A produção de derivados de petróleo requer um controle rigoroso dos parâmetros de qualidade. O combustível utilizado em aviões, por exemplo, tem que permanecer no estado líquido, mesmo exposto a baixas temperaturas. Para verificar se esta condição é atendida são realizados ensaios do ponto de congelamento<sup>1</sup>.

A maioria dos processos industriais é bem equipada com sensores de variáveis de processo (temperatura, vazão, pH, etc.). A medição destas variáveis é feita através dos sensores físicos, como os de temperatura, os de vazão, os de pressão, etc. Já os parâmetros de qualidade (concentração, por exemplo) não contam com uma boa oferta de sensores nas plantas industriais. A medição destes parâmetros é feita através dos sensores químicos ou analíticos, que são dispositivos capazes de determinar a concentração de componentes químicos em líquidos ou gases, e convertem esta informação em um sinal elétrico ou ótico. Todos estes sensores (físicos ou analíticos) disponibilizam dados em tempo real.

Para os parâmetros de qualidade, geralmente não existem sensores disponíveis no mercado, ou os sensores são caros. Como alternativa são utilizados os ensaios de laboratório, que apresentam o inconveniente de levar um longo tempo, de duas a doze horas, para os resultados das análises estarem disponíveis. Como consequência desta demora, por precaução, adota-se uma produção conservativa<sup>2</sup> de derivados. Ainda assim pode ocorrer produção fora de especificação, acarretando reprocessamento.

Com a finalidade de garantir o conhecimento dos parâmetros de qualidade destes derivados em menor tempo, utilizam-se sensores virtuais, que são programas de computador que calculam estes parâmetros, através de dados de medições em tempo real. O que fundamenta a utilização destes sensores virtuais é o fato dos parâmetros de qualidade dos derivados serem dependentes de variáveis do processo, tais como vazão, temperatura e pressão.

Uma vez que a relação entre os parâmetros de qualidade e as variáveis de processo geralmente é não linear, a abordagem de rede neuronal é uma opção conveniente para a sua modelagem. Além disso, é razoável supor que há uma dependência contínua dos parâmetros de qualidade com relação às variáveis de processo. Deste modo podemos explorar, também, a capacidade da rede neuronal de aproximar qualquer função contínua.

As colunas de destilação, que são responsáveis por separar uma mistura em duas ou mais outras misturas, não conseguem realizar esta separação de forma ideal. Nas refina-

---

<sup>1</sup> Ponto de congelamento: temperatura em que se inicia a mudança de fase da substância de líquida para sólida.

<sup>2</sup> Conservadora, com margem de segurança, abaixo da produção ótima.

rias, complexos industriais que produzem derivados de petróleo, a torre depropanizadora é uma coluna que tem o objetivo de separar butano ( $C_4$ ) do propano ( $C_3$ ). O propano, que sai pelo topo, sempre conterà frações pesadas, o mesmo ocorre com o butano, produto do fundo, em relação às frações leves.

Estas colunas possuem fluxos frios e quentes que possibilitam que ajustes sejam feitos nos parâmetros de qualidade. Além disso, medições de temperaturas e pressões indicam, de forma indireta, em que valores estão, aproximadamente, estes parâmetros de qualidade. Estas medições foram utilizadas como entradas de uma rede neuronal.

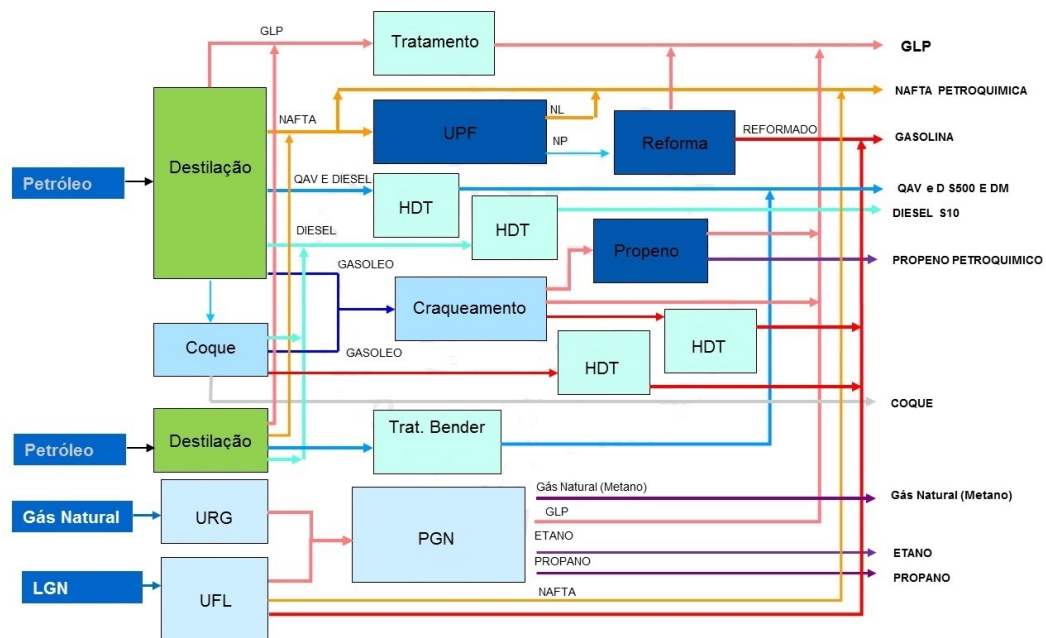
## 1 DERIVADOS DE PETRÓLEO

O petróleo é uma mistura homogênea de compostos orgânicos, principalmente de hidrocarbonetos, moléculas formadas por átomos de hidrogênio e carbono. É uma substância oleosa e inflamável sendo a principal fonte de energia da atualidade. Através de diferentes processos é possível produzir diversos derivados, como nafta petroquímica, gasolina, querosene de aviação, diesel e gás liquefeito de petróleo (GLP)<sup>3</sup>.

A cadeia produtiva de uma refinaria é composta por várias unidades de processo que atuam em sequência, como pode ser visto na Figura 1. Estas unidades ou realizam processos físicos, onde não ocorre mudança das substâncias presentes na carga<sup>4</sup>, ou processos químicos, onde ocorre quebra de moléculas, modificando as substâncias presentes originalmente na carga.

A cadeia começa nas unidades de destilação que recebem o petróleo e, através de processos físicos, separam-no em diversos fluxos que seguirão como carga de outras unidades. Há também unidades dedicadas exclusivamente ao tratamento do derivado, que removem contaminantes que comprometem a utilização do mesmo.

Figura 1 - Esquema de Produção de Derivados de uma Refinaria.



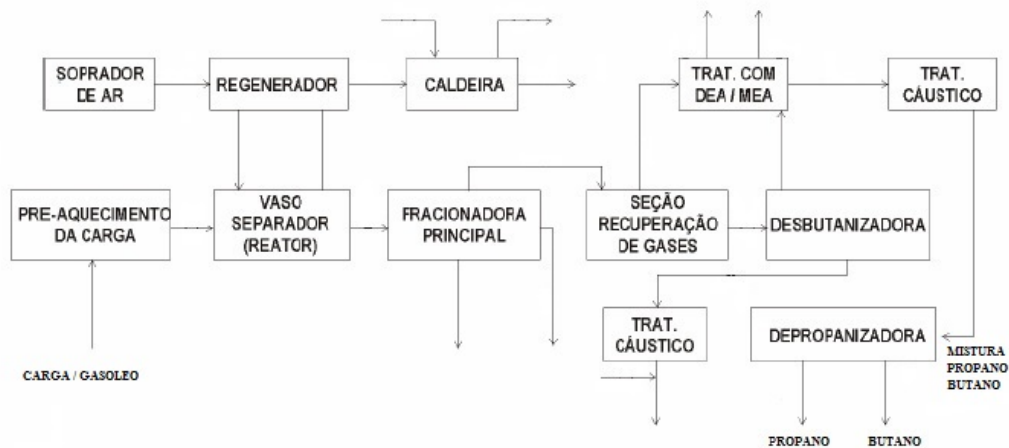
Fonte: CAMARGO, 2001.

<sup>3</sup> Gás presente nos botijões utilizados em fogões domésticos.

<sup>4</sup> Denomina-se por carga o fluxo de entrada, ou alimentação, de unidades de processo e de determinados equipamentos industriais de uma refinaria.

Uma das unidades de maior rentabilidade de uma refinaria é a de craqueamento catalítico. Uma representação esquemática, das varias fases do processo desta unidade, é vista na Figura 2. Esta unidade, de processo químico, recebe a carga das unidades de destilação e produz GLP, gasolina e cargas para outras unidades. A coluna depropanizadora é o último estágio nesta unidade para a produção de GLP, e será descrita na próxima seção.

Figura 2 - Esquema de uma Unidade de Craqueamento Catalítico.



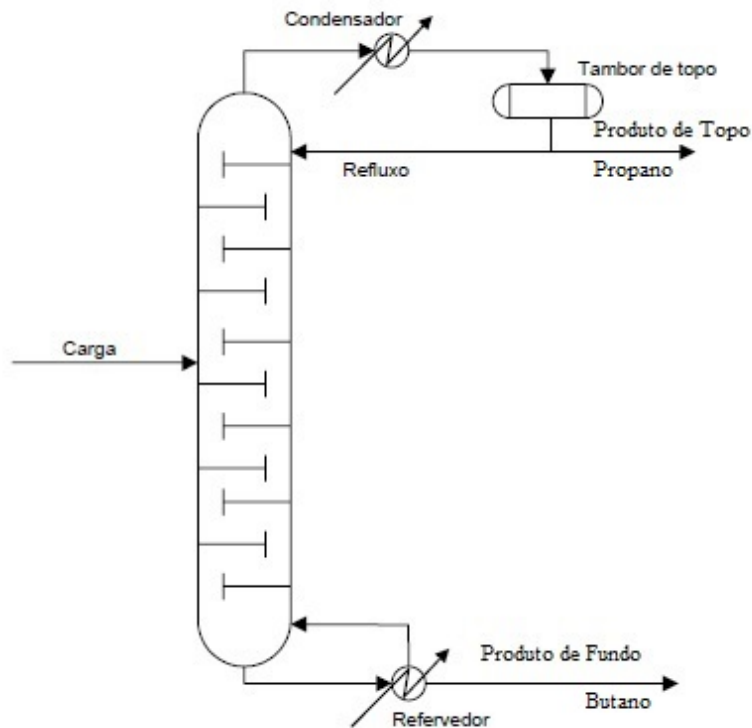
Fonte: O autor, 2016.

## 1.1 Coluna Depropanizadora

A coluna depropanizadora recebe como carga uma mistura de propano e butano. Sua função é separá-la, por meio do processo de destilação, em duas vazões. O propano, produção de topo, é enviado como carga para outra unidade e o butano, produção do fundo, fará parte da produção de GLP da refinaria. A carga inicial é aquecida antes de entrar na coluna, desta forma, parte dela vaporiza, sendo que o vapor sobe através da coluna e o líquido vai para o fundo.

No topo da coluna, há a vazão de refluxo, que é parte do fluxo de topo, que após ser resfriado e condensado, retorna à coluna, condensando as frações mais pesadas do vapor que sobe. Da mesma forma, parte do fluxo de fundo é aquecida e retorna à coluna, fazendo com que as frações mais leves do fundo vaporizem (Figura 3).

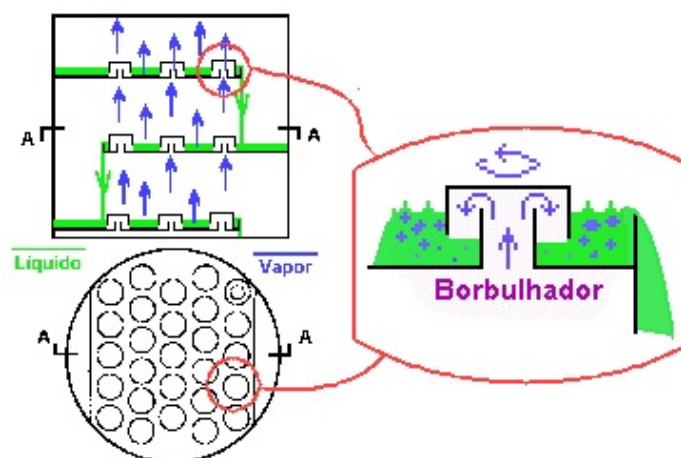
Figura 3 - Coluna Depropanizadora.



Fonte: CAMARGO, 2001.

A coluna depropanizadora possui dezenas de pratos com borbulhadores (Figura 4), que direcionam os fluxos de vapor e de líquido, possibilitando o contato entre o vapor que sobe pela coluna e o líquido que desce. Este contato tem o objetivo de purificar os dois fluxos, desta forma, reduzindo o percentual de frações pesadas (butano) na vazão de topo e o percentual de frações leves (propano) na vazão de fundo.

Figura 4 - Prato com borbulhadores.



Fonte: USP, 2016.

A separação dos produtos na coluna não é ideal, por isso, durante o processo

de separação, a vazão de topo contém frações pesadas que deveriam fazer parte da vazão de fundo, assim como a de fundo contém frações leves da vazão de topo.

Em função da utilização futura desses produtos, estas vazões possuem especificações de qualidade que determinam os teores máximos de frações pesadas na vazão de topo (propano) e de leves na vazão de fundo (butano). Por questões de utilização, o teor máximo de frações mais pesadas que o propano ( $C_3$ ) no produto de topo é de 4%. E por questões econômicas, o teor de frações mais leves que o butano ( $C_4$ ) no produto de fundo deve ser o menor possível.

Uma das razões para a produção ficar fora de especificação é a alteração na composição da carga da coluna. Esta alteração pode ser devida a uma troca de tanque ou ajustes na unidade de craqueamento catalítico. Os tanques que armazenam o petróleo levam de dois a três dias para serem processados nas unidades de destilação. A mudança de tanque pode acarretar a ocorrência de diferenças na composição do petróleo, que se não forem perceptíveis, permitindo ajustes no processo, poderão comprometer parâmetros de qualidade de algum derivado.

Este problema só será identificado e corrigido após os resultados dos ensaios de laboratórios estarem disponíveis. Ajustes feitos no processo da unidade de craqueamento catalítico, por exemplo, na fracionadora principal, poderão alterar a composição da carga da coluna depropanizadora. Tal como na mudança de tanque na unidade de destilação, esta alteração na composição poderá comprometer parâmetros de qualidade do butano e/ou propano.

A garantia de produção de propano e butano com frações pesadas e leves dentro dos limites exigidos é de grande interesse econômico, quer seja pelo fato da produção estar próxima da ótima ou por evitar reprocessamento. Os necessários ajustes da coluna, para garantir a produção dentro das especificações, exigem o conhecimento dos valores destas frações, que podem ser vistas como contaminantes.

Como já foi dito, o controle por análise de laboratório não é eficaz em função do longo tempo necessário para os resultados estarem disponíveis. Por isso existe um grande esforço em desenvolver modelos que calculem estas especificações a partir das variáveis de processo associadas. Os bancos de dados de processo das refinarias, que possuem um volume muito grande de informação, permitem que as redes neurais sejam utilizadas para gerarem estes modelos.

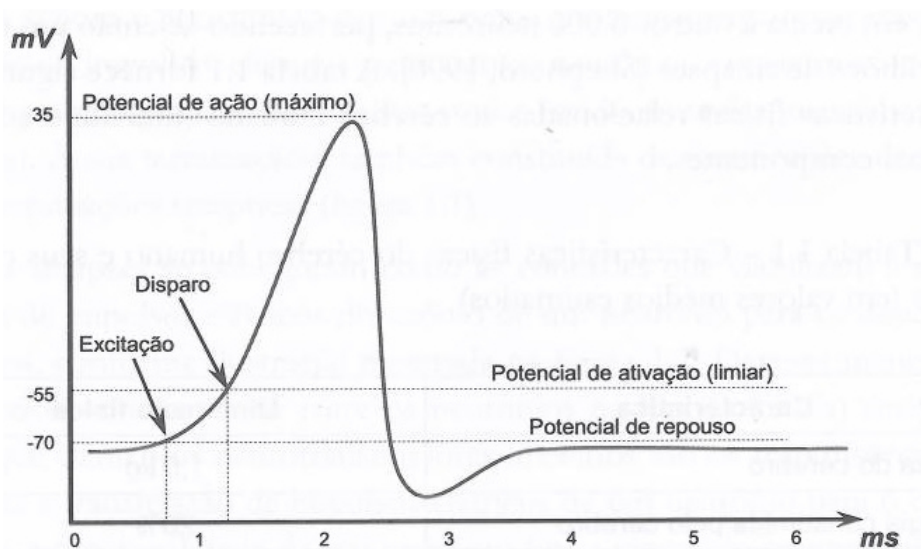
## 2 REDE NEURONAL ARTIFICIAL

O cérebro humano é constituído por elementos biológicos denominados neurônios, operando em paralelo, produzindo ações adequadas para cada uma das suas funcionalidades cerebrais e mentais, tais como memorizar e raciocinar (SILVA; SPATTI; FLAUZINO, 2010). Redes neuronais artificiais são programas de computador inspirados no mecanismo do cérebro dos seres humanos.

### 2.1 Neurônio Biológico

A unidade elementar do sistema cerebral é o neurônio e sua função é transmitir impulsos elétricos em determinadas condições de operação (SILVA; SPATTI; FLAUZINO, 2010). As sinapses são conexões responsáveis pelas transmissões destes impulsos de um neurônio para outro. Não existe contato físico entre as membranas das terminações neuronais das conexões sinápticas. Os elementos neurotransmissores, liberados na junção sináptica, são encarregados por ponderar a transmissão de impulsos elétricos de um neurônio para outro. Na verdade, a funcionalidade de um neurônio é dependente da atuação dessas conexões sinápticas que são dinâmicas e dependentes da química cerebral (HODGKIN; HUXLEY, 1952). Quando o neurônio é estimulado além de seu limiar de ativação, ele dispara um impulso elétrico, que será propagado para os demais neurônios a ele interligados. A Figura 5 apresenta as fases envolvidas nas mudanças de potencial de ação no neurônio ao longo de sua excitação.

Figura 5 - Mudanças de potencial do neurônio durante sua excitação.



Fonte: SILVA; SPATTI; FLAUZINO, 2010.

Cientistas brasileiros, liderados pela neurocientista Suzana Herculano-Houzel, concluíram que o cérebro humano possui aproximadamente oitenta e seis bilhões de neurônios<sup>5</sup>, e não cem bilhões como se pensava anteriormente, que interligados dão ao indivíduo a capacidade de inteligência e raciocínio, memória e emoções, a motricidade, o equilíbrio, a postura corporal, a percepção e reconhecimento de estímulos auditivos e visuais. O cérebro controla os batimentos cardíacos, o balanço hídrico, a temperatura corporal e a pressão arterial. As redes neuronais artificiais tentam simular o funcionamento deste “equipamento” extraordinário que é o cérebro humano.

## 2.2 Neurônio Artificial

Uma rede neuronal artificial pode ser expressa como um conjunto de unidades de processamento, chamadas de neurônios artificiais, que são interligados por um grande número de interconexões (sinapses artificiais), representadas pelos pesos sinápticos. As redes neuronais artificiais, como o cérebro humano, possuem capacidade de aprendizagem, de adaptação por experiência e de generalização.

O neurônio artificial é o responsável pelo processamento da informação para a execução de uma rede neuronal artificial. Os pesos sinápticos, que correspondem às sinapses do neurônio biológico, são valores que ponderam os sinais de entrada de cada neurônio, oriundos do meio externo ou associados às saídas de outros neurônios. A entrada  $x_j$  transmitida através da sinapse que conecta o neurônio  $j$  ao neurônio  $i$  é multiplicada pelo peso sináptico  $w_{ij}$ , que quantifica a relevância desta entrada em relação à funcionalidade do respectivo neurônio destino, neurônio  $i$ . (SILVA; SPATTI; FLAUZINO, 2010).

Os neurônios de uma rede neuronal biológica podem estar nos níveis excitatórios ou inibitórios. No neurônio artificial o nível inibitório é representado, nos modelos matemáticos, por conexão sináptica negativa. Na rede neuronal, o neurônio  $i$  recebe um estímulo  $\vartheta_i$ . Ele reage a este estímulo e apresenta a resposta  $y_i$ . O estímulo  $\vartheta_i$  é a ponderação pelos pesos sinápticos dos sinais de entrada do neurônio  $i$ .

## 2.3 Função de Ativação

A função de ativação é responsável por transformar o nível de ativação (estímulo) de um neurônio em seu sinal de saída. É desejável que a função de ativação seja não linear, continuamente diferenciável, limitada, suave e monótona. A função ser di-

---

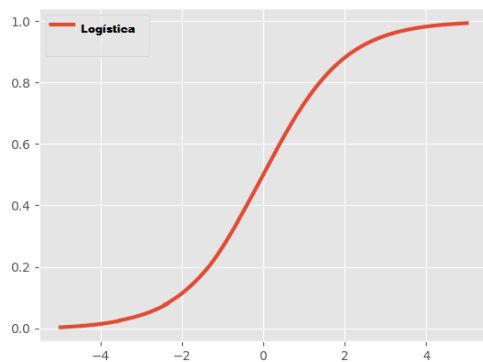
<sup>5</sup> [http://www2.uol.com.br/vivermente/noticias/recontando\\_neuronios.html](http://www2.uol.com.br/vivermente/noticias/recontando_neuronios.html), acessado em 22/09/2016.



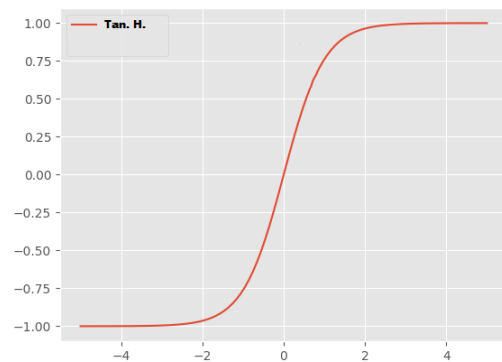
ferenciável é um requisito do algoritmo *backpropagation* (retropropagação de erros), como será visto mais adiante. Uma função limitada tende a tornar o método do gradiente, que faz parte do algoritmo de retropropagação, mais estável.

As funções de ativação mais utilizadas, nas redes do tipo perceptron de múltiplas camadas, são a função logística e a tangente hiperbólica. A função logística (Figura 6a) tem sua imagem num intervalo fechado contínuo de  $[0, 1]$ . Pode-se estender este intervalo para  $[-1, 1]$  utilizando a tangente hiperbólica (Figura 6b).

Figura 6 - Função de Ativação.



(a)



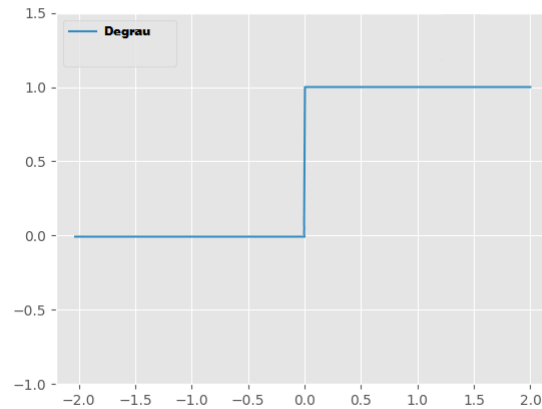
(b)

Legenda: (a) – Função logística; (b) – Função tangente hiperbólica.

Fonte:FACURE, 2017.

A escolha da função de ativação pode influenciar fortemente o desempenho e a complexidade da rede. Artigos mostram que esta escolha é considerada, por muitos especialistas, tão importante como a arquitetura e o algoritmo de aprendizagem da rede (GOMES, 2010). O uso do algoritmo de retropropagação de erros exige que a função de ativação seja suave, isto é, diferenciável, diferentemente da função degrau (Figura 7) usada pelo perceptron original.

Figura 7 - Função degrau.



Fonte: FACURE, 2017.

Segundo Skoundrianos e Tzafestas (1991), a forma da função de ativação e de sua derivada pode afetar o resultado do treinamento. Wen e Ma (2005) propuseram uma rede neural Max-Piecewise-Linear (MPWL) para aproximação de funções. A rede MPWL consiste de uma única camada oculta e emprega as funções de base Piecewise-Linear (PWL) como as funções de ativação dos neurônios ocultos. É possível encontrar aplicações em uma variedade de campos como síntese do circuito não linear, identificação de sistemas dinâmicos e controle.

### 2.3.1 Saída do Neurônio

Em termos matemáticos, a saída  $y_i$  do neurônio  $i$  é obtida através de

$$y_i = \varphi(\vartheta_i), \quad (1)$$

onde  $\varphi$  é função de ativação do neurônio  $i$ , e o estímulo  $\vartheta_i$  é dado por

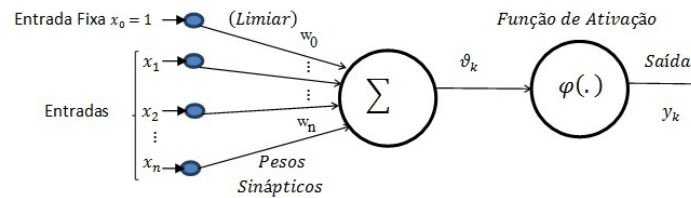
$$\vartheta_i = \sum_{j=1}^n (x_j * w_{ij}) - b_i, \quad (2)$$

isto é, a soma ponderada das entradas  $x_j$  pelos respectivos pesos sinápticos  $w_{ij}$  comparada com o limiar de ativação  $b_i$ .

Na implementação computacional, o limiar  $b_i$  é representado pela entrada fixa  $x_0 = +1$ , e pelo peso sináptico  $w_0 = b_i$ . A operação soma representada na Figura 8, que representa o modelo do neurônio artificial, efetua uma combinação linear dos sinais de entrada ponderados pelas respectivas sinapses do neurônio, resultando no sinal  $\vartheta_i$ . O “potencial de ativação” é o resultado de aplicar  $\vartheta_i$  à “função de ativação”,  $\varphi$ , ou seja, é

o sinal de saída do neurônio,  $y_i$ . Se esta diferença é positiva o neurônio produz um sinal excitatório, se não, será inibitório.

Figura 8 - Modelo de um neurônio artificial.



Fonte: O autor, 2016.

Considerando o neurônio da Figura 8, as entradas  $x_i = 1, \dots, n$  são oriundas do mundo externo (entradas da rede neuronal) ou das saídas dos neurônios de camadas anteriores ou posteriores<sup>6</sup>. Se o limiar é incorporado ao somatório em (2), podemos redefinir o estímulo e a resposta no neurônio  $i$  da seguinte maneira

$$\vartheta_i = \sum_{j=0}^n (x_j * w_{ij}), \quad (3)$$

e

$$y_i = \varphi(\vartheta_i). \quad (4)$$

As equações (3) e (4) representam o modelo matemático-computacional do neurônio.

## 2.4 Redes Neurais Artificiais

A área de estudos de redes neuronais artificiais envolve várias disciplinas: matemática, neurociência, computação, estatística e engenharia. Redes neuronais artificiais podem ser aplicadas em problemas de aproximação funcional, análise de séries temporais, reconhecimento de padrões, processamento de sinais, identificação e controle de processos, em função de uma importante propriedade: “a habilidade de apreender com os dados” (HAYKIN, 2003).

Trabalhos teóricos (CYBENKO, 1989; FUNAHASHI, 1989) provaram que, com apenas uma camada oculta, uma rede neuronal pode aproximar qualquer função contínua. Esse resultado estabelece uma base teórica sólida para a modelagem de sistemas não lineares com redes neuronais artificiais.

---

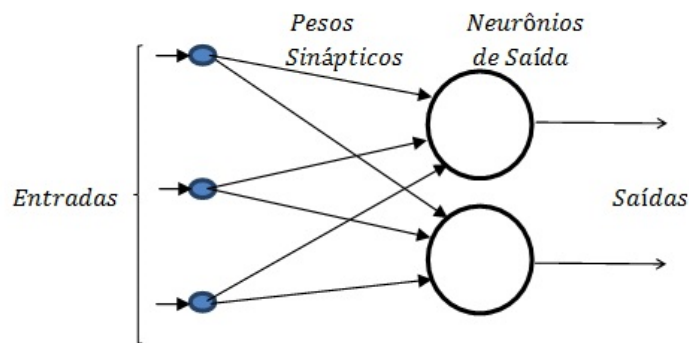
<sup>6</sup> Redes recorrentes, que serão vistas mais adiante.

A arquitetura de uma rede neuronal artificial define a forma como os seus diversos neurônios estão arranjados, ou dispostos uns em relações aos outros (SILVA; SPATTI; FLAUZINO, 2010). A topologia define o número de camadas e de neurônios de cada uma destas camadas.

Destá forma, em relação à arquitetura, as redes neuronais estão classificadas em:

- Perceptron Simples (Figura 9) e Adaline são redes de alimentação à frente de camadas simples. Seu algoritmo de treinamento é geralmente especificado como regra delta, ou regra do mínimo quadrático médio, ou regra de Widrow-Hoff ou regra adaline (HERTZ; PALMER; KROGH, 1991).

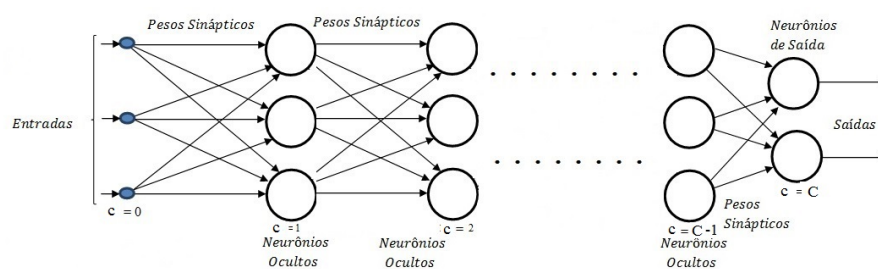
Figura 9 - Perceptron simples.



Fonte: O autor, 2016.

- Redes perceptron de múltiplas camadas (Figura 10) e redes de base radial são redes de alimentação à frente de camadas múltiplas. A diferença entre elas está no algoritmo de treinamento, regra delta generalizada ou retropropagação de erros e regra delta competitiva, respectivamente. Estas redes possuem camadas ocultas que são caracterizadas por não se comunicarem com o meio externo.

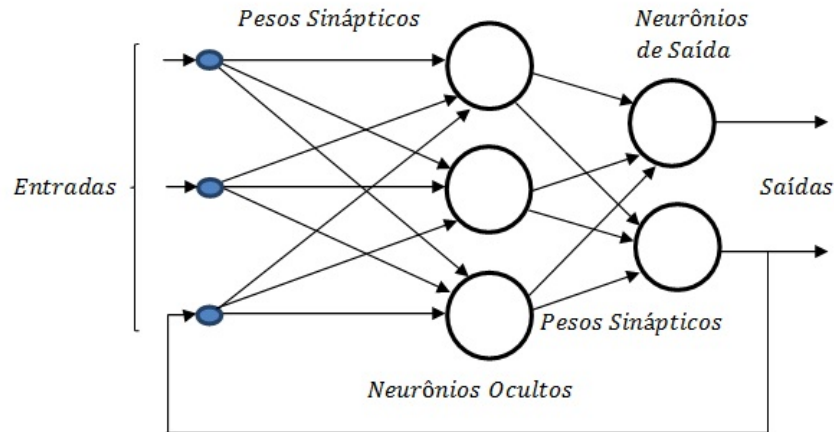
Figura 10 - Perceptron de múltiplas camadas.



Fonte: O autor, 2016.

- Redes recorrentes e de Hopfield possuem pelo menos um laço de realimentação (Figura 11).

Figura 11 - Rede recorrente.



Fonte: O autor, 2016.

A capacidade de aprendizado da rede foi teoricamente estabelecida por Cybenko (1989). Ele provou que toda função real de  $n$  variáveis reais contínuas no hipercubo unitário pode ser uniformemente aproximada por uma rede do tipo perceptron de múltiplas camadas, com uma camada oculta. Para provar seu resultado, Cybenko considerou funções discriminatórias definidas no hipercubo unitário  $I_n = [0, 1]^n$ . Estas funções  $\sigma$  definidas em  $I_n$  tais que se, para uma medida de Borel regular sobre  $I_n$ , finita e com sinal, temos

$$\int_{I_n} \sigma(y^t + \theta) d\mu(x) = 0,$$

para todo  $y \in \mathbb{R}^n$  e  $\theta \in \mathbb{R}$  então  $\mu = 0$ .

**Teorema 1** *Seja  $\sigma$  uma função de ativação discriminatória e contínua. Então combinações lineares finitas da forma*

$$G(x) = \sum_{j=1}^N \alpha_j \sigma(y_j^t x + \theta_j),$$

com  $N \in \mathbb{N}$ ,  $\alpha_j, \theta_j \in \mathbb{R}$  e  $x, y_j \in \mathbb{R}^n$  para  $j = 1, \dots, N$ , são densas no espaço  $C([0, 1]^n)$  das funções contínuas no hipercubo unitário.

Nos experimentos a função de ativação utilizada foi a tangente hiperbólica que, de acordo com a definição, é discriminatória.

## 2.5 Retropropagação do Erro

A fim de que a rede seja capaz “de aprender com os dados”, ela precisa ser treinada. O algoritmo de treinamento, retropropagação de erros, é um dos mais utilizados para treinamento de redes do tipo perceptron de múltiplas camadas.

Ele consiste em um algoritmo supervisionado que utiliza pares (entrada, saída desejada) para, por meio de um mecanismo de correção de erros, ajustar os pesos da rede. O treinamento ocorre em duas fases, cada uma delas percorre a rede em um sentido: para frente (*forward*) em que a saída da rede é definida para determinado padrão de entrada, e para trás (*backward*), em que se utiliza uma propagação do erro (diferença entre as saídas desejada e produzida pela rede) da camada de saída para as camadas ocultas, de forma a atualizar os pesos das conexões (PADUA et al., 2000).

Basicamente, o algoritmo de treinamento de forma supervisionada, na rede perceptron de múltiplas camadas, consiste de:

- Sinais de entrada que se propagam para frente, camada por camada, gerando sinais de saída.
- Diferenças entre os sinais de saída e as saídas desejadas que geram os sinais de erros da rede.
- Sinais de erros se propagam de volta na rede, em sentido contrário aos de entrada e que servem para ajustar os pesos, para que as saídas da rede se aproximem dos valores desejados.

## 2.6 Algoritmo de Retropropagação de Erros - Backpropagation

A seguir uma breve descrição do algoritmo retropropagação de erros (HERTZ; PALMER; KROGH, 1991). Onde:

- $\xi_j^\mu$  representa  $j$ -ésima entrada da rede da iteração  $\mu$ .
- $C$  representa o número total de camadas, ocultas e a de saída, e  $c$  uma camada específica.
- $w_{ij}^c$  é o peso sináptico conectando o  $j$ -ésimo neurônio da camada  $c - 1$  ao  $i$ -ésimo neurônio da camada  $c$ .
- $\vartheta_i^c$  é a entrada ponderada em relação ao  $i$ -ésimo neurônio da camada  $c$  conforme Equação (3).

- $\varphi$  é a função de ativação de cada neurônio.
  - $y_i^c$  é a saída do  $i$ -ésimo neurônio da camada  $c$ .
1. Inicialize os pesos sinápticos com valores aleatórios pequenos no intervalo  $[-1, 1]$ .
  2. Escolha um padrão  $\xi_k^\mu$  e o aplique à camada de entrada ( $c = 0$ ) de forma que

$$y_k^0 = \xi_k^\mu,$$

para todo  $k$ .

3. Propague o sinal adiante através da rede usando

$$y_i^c = \varphi(\vartheta_i^c) = \varphi\left(\sum_j w_{ij}^c y_j^{c-1}\right),$$

para cada  $i$  e  $c$  até que as saídas finais  $y_i^C$  tenham todas sido calculadas.

4. Calcule os deltas para a camada de saída de acordo com

$$\delta_i^C = \varphi'(\vartheta_i^C) [\zeta_i^\mu - y_i^C],$$

comparando as saídas atuais  $y_i^C$  com os valores desejados  $\zeta_i^\mu$  para o padrão  $\mu$  que está sendo considerado.

5. Calcule os deltas para as camadas anteriores propagando o erro para trás usando

$$\delta_i^{c-1} = \varphi'(\vartheta_i^{c-1}) \sum_j w_{ji}^c \delta_j^c,$$

para  $c = C, C - 1, \dots, 2$  até que um delta tenha sido calculado para cada neurônio.

6. Use

$$\Delta w_{ij}^c = \eta \delta_i^c y_j^{c-1},$$

para atualizar todos os pesos sinápticos de acordo com

$$w_{ij}^{atual} = w_{ij}^{anterior} + \Delta w_{ij}.$$

7. Volte ao item (2) e repita para o próximo padrão.

### 3 EXPERIMENTOS PRELIMINARES

Com o objetivo de aprofundar o conhecimento sobre o treinamento de uma rede neuronal para o cálculo de uma função contínua e a escolha da topologia mais adequada, foram realizados experimentos com uma função conhecida. A arquitetura escolhida para os testes foi a do perceptron de múltiplas camadas, que vem sendo utilizada com sucesso para resolver diversos problemas complexos, através do treinamento de forma supervisionada, com o algoritmo de retropropagação de erros. As funções escolhidas para serem aprendidas pela rede foram  $\sin x$  e o  $\cos x$ .

Um experimento é composto por cinco treinamentos, conforme passos apresentados nas seções 2.5 e 2.6. Em cada treinamento é executado um ciclo completo do algoritmo e o mesmo conjunto de amostras é utilizado. O que diferencia cada treinamento é o conjunto de pesos sinápticos iniciais, que são distintos para cada um. Um treinamento é composto por épocas, onde é apresentado o mesmo conjunto completo de amostras, mas a ordem de apresentação é diferente em cada época, sendo esta ordem de apresentação determinada por uma ordenação aleatória.

Após o término dos treinamentos, o erro da fase de treinamento é calculado de acordo com a Equação (5), utilizando o conjunto de dados da fase de treinamento e o conjunto  $W$  de pesos sinápticos que apresentou melhor desempenho nos treinamentos. Logo após, este mesmo conjunto  $W$  será utilizado na fase de teste, para calcular o erro desta fase, novamente utilizando a Equação (5) e o conjunto de dados da fase de teste, separados conforme critério descrito abaixo.

O conjunto de dados foi dividido, aleatoriamente, em 70% para o conjunto da fase de treinamento e 30% para o da fase de teste.

#### 3.1 Topologia

Como escolhemos funções reais de uma variável, ficou definido uma unidade na camada de entrada e outra na camada de saída. O número de camadas ocultas e o de neurônios, em cada uma delas, serão definidos com vistas a obter menor erro quadrático médio possível.

Para cálculo do erro considera-se a diferença entre a saída calculada pela rede e o valor conhecido da função seno ou cosseno, para o valor de entrada da variável  $x$



apresentada à rede. Precisamente, o erro quadrático médio  $E_m$  é dado por

$$E_m = \sqrt{\frac{1}{p} \sum_{k=1}^p (\zeta_k - y_k)^2}, \quad (5)$$

onde  $p$  é o número de amostras,  $\zeta_k$  é o valor conhecido da função (valor desejado) e  $y_k$  é a saída produzida pela rede, ambos associados ao valor de entrada  $x_k$ . Nos testes, como função de ativação dos neurônios das camadas ocultas, foi utilizada a tangente hiperbólica e da camada de saída a função identidade.

### 3.2 Derivada da Função de Ativação

O algoritmo de retropropagação de erros requer o conhecimento da derivada da função de ativação, o que, para a função  $\sigma(x) = \tanh(x)$ , é facilmente deduzida

$$\frac{d\sigma}{dx} = \text{sech}^2 x = (1 - \tanh^2 x) = (1 - \tanh x)(1 + \tanh x) = (1 - \sigma)(1 + \sigma). \quad (6)$$

Durante a implementação numérica do algoritmo de retropropagação, o valor da derivada da tangente hiperbólica é armazenado no computador em uma variável de ponto flutuante. Para destacar a diferença entre o valor teórico exato e o representado no computador, denotaremos com um asterisco as variáveis de ponto flutuante, por exemplo,  $\sigma^*$  é a versão em ponto flutuante de  $\sigma$ . Desse modo, no computador a derivada da tangente hiperbólica é dada por

$$\left(\frac{\partial\sigma}{\partial x}\right)^* = (1 - \sigma^*)(1 + \sigma^*). \quad (7)$$

Mas esta expressão está sujeita a erros de arredondamentos. Se o valor de  $\sigma$  é próximo de um, então  $\sigma^*$  é exatamente um, desta forma  $\left(\frac{\partial\sigma}{\partial x}\right)^*$  será zero (PLATE, 1988). A derivada zero pode estacionar o processo de treinamento.

### 3.3 Teste inicial

O intervalo  $\mathcal{I} = [-15.708, 31.41]$  foi fixado, como domínio da função a ser aprendida pela rede. Consideramos uma partição  $\mathcal{P} = \{-15.708 = x_1 < x_2 < \dots < x_{500} = 31.41\}$  do intervalo  $\mathcal{I}$ , com 500 pontos igualmente espaçados, com intervalo de 0.094, isto é,  $\mathcal{P} = \{x_1 < x_2 < \dots < x_{500}\}$  com  $x_1 = -15.708$  e  $x_j = x_1 + (j - 1) \cdot 0.094$

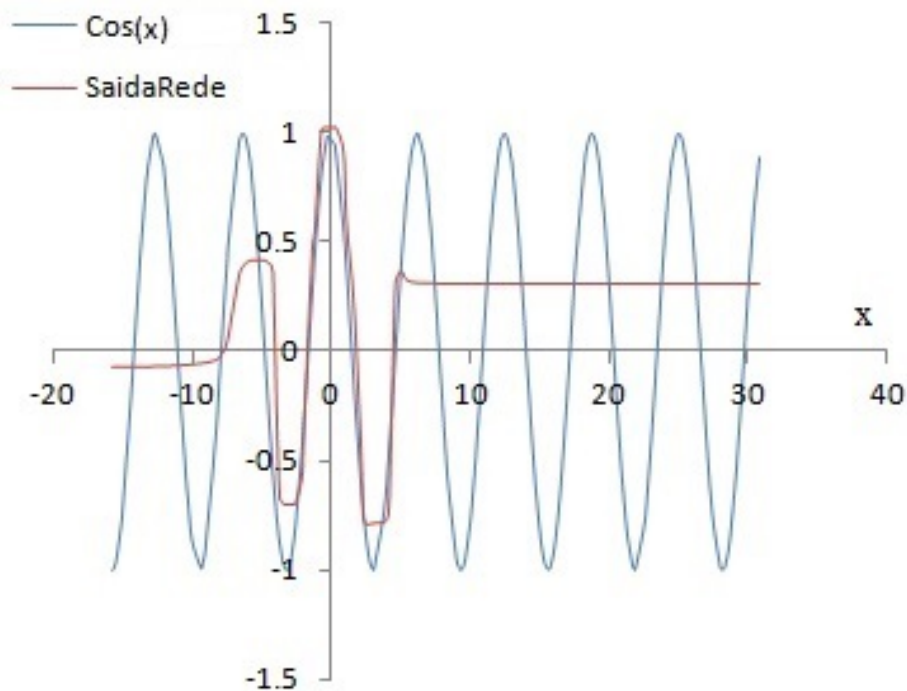
para  $j \in \mathcal{J} = \{1, \dots, 500\}$ . Conforme mencionado anteriormente, 70% dos pontos de  $\mathcal{P}$  foram selecionados para gerar os pares ordenados (entrada, saída desejada) do conjunto da fase de treinamento da rede e os 30% restantes foram utilizados para o conjunto da fase de teste.

Os critérios de parada da fase de treinamento foram:

- Número máximo de épocas igual a 10000.
- Erro quadrático médio menor que 0.01, considerando que este erro já produz uma boa aproximação.
- Estabilização do erro quadrático médio após cinco épocas consecutivas (diferença absoluta entre épocas menor que 0.1).

Apesar de ter sido provado que uma rede, com apenas uma camada oculta, seja capaz de aproximar qualquer função contínua, depois de várias topologias testadas, o melhor ajuste da rede que conseguimos realizar foi o de uma rede com quatro camadas ocultas com 30, 20, 10 e 5 neurônios e esta não obteve êxito para aproximar a função cosseno, como pode ser visto na Figura 12.

Figura 12 - Cosseno e a saída produzida pela rede neuronal sem convergir com topologia de uma unidade na camada de entrada, 30 neurônios na primeira camada oculta, 20 neurônios na segunda camada, 10 na terceira, 5 na quarta e um neurônio na camada de saída.



Fonte: O autor, 2016.

Foi constatado o problema de saturação da função de ativação, para valores absolutos de entrada,  $x$ , maiores que 6.5. Esse problema foi solucionado, em parte, pela

normalização dos valores de entrada de acordo com a Equação (8), que produz como resultado final valores no intervalo  $[-1, 1]$ , que corresponde a imagem da função tangente hiperbólica.

$$Entrada_{normalizada} = \frac{2 * (x - x_{min})}{x_{max} - x_{min}} - 1 \quad (8)$$

### 3.4 Resultados

Cada experimento foi composto de cinco treinamentos, e em cada um o número máximo de épocas foi de dez mil. Os experimentos começaram com uma topologia bem simples, uma camada oculta com dois neurônios e com o intervalo da variável  $x$  em  $[-5.0, 5.0]$ , já que a saturação aconteceu com o intervalo de  $[-15.708, 31.41]$ . Conforme o erro quadrático médio estabilizava, aumentava-se o número de neurônios da camada oculta. Quando esta estratégia se tornava ineficiente, isto é, não ocorria redução do erro médio, significando que com a topologia atual não haveria melhora do aprendizado da rede, aumentava-se o número de camadas ocultas.

O treinamento foi feito com o algoritmo de retropropagação de erros. Em cada experimento, inicialmente foi fixado um intervalo fechado e limitado  $\mathcal{X} = [a, b]$ . Foi considerada uma partição  $\mathcal{I}$  do intervalo  $\mathcal{P}$ , com pontos igualmente espaçados, isto é,  $\mathcal{P} = \{a < x_2 < \dots < x_t = b\}$  com  $x_1 = a$  e  $x_j = x_1 + \frac{(j-1)}{(b-a)}(t-1)$  para  $j = 1, \dots, t$ .

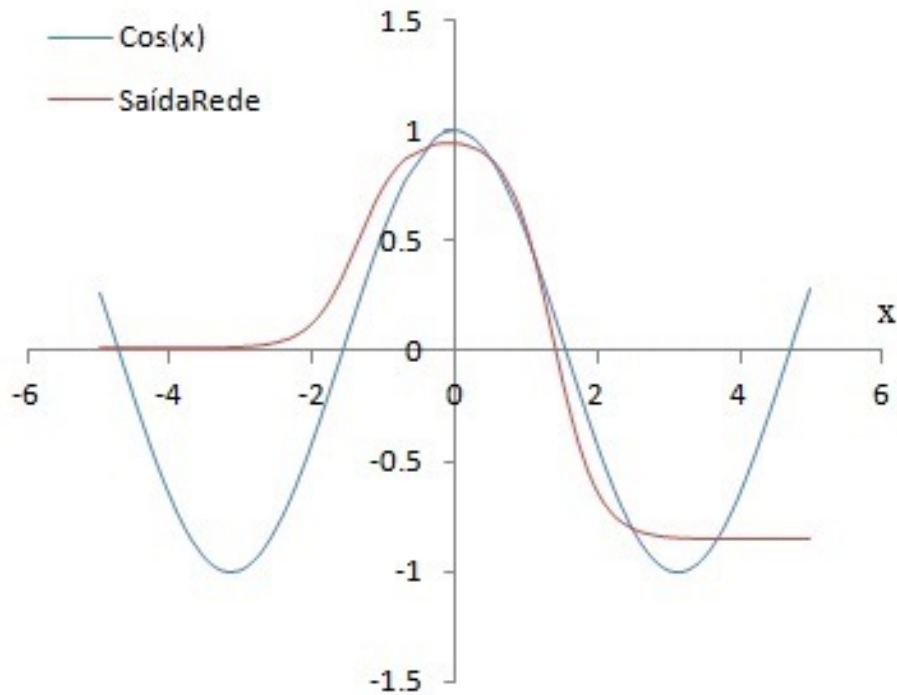
Na descrição dos resultados, indicaremos a topologia da rede, o intervalo  $\mathcal{I}$  e o critério de parada que foi atendido durante o treinamento. Como mencionado anteriormente, em todos os experimentos, a rede tinha uma unidade na camada de entrada, assim como, um neurônio na de saída. A seguir são apresentados os resultados obtidos. Os erros apresentados neste primeiro experimentos são da fase de treinamento.

#### 3.4.1 Uma camada oculta.

##### 3.4.1.1 Uma camada oculta com dois neurônios.

Neste experimento, o intervalo considerado da variável  $x$  foi de  $\mathcal{I} = [-5.0, 5.0]$  e o critério de parada atendido foi que o número máximo de épocas foi atingido. O menor erro, de acordo com a Equação (5), foi de 0.325617. A rede não fez uma boa aproximação, como pode ser visto na Figura 13.

Figura 13 - Cosseno e a saída da rede com topologia de uma unidade na camada de entrada, 2 neurônios na camada oculta e um neurônio na camada de saída.

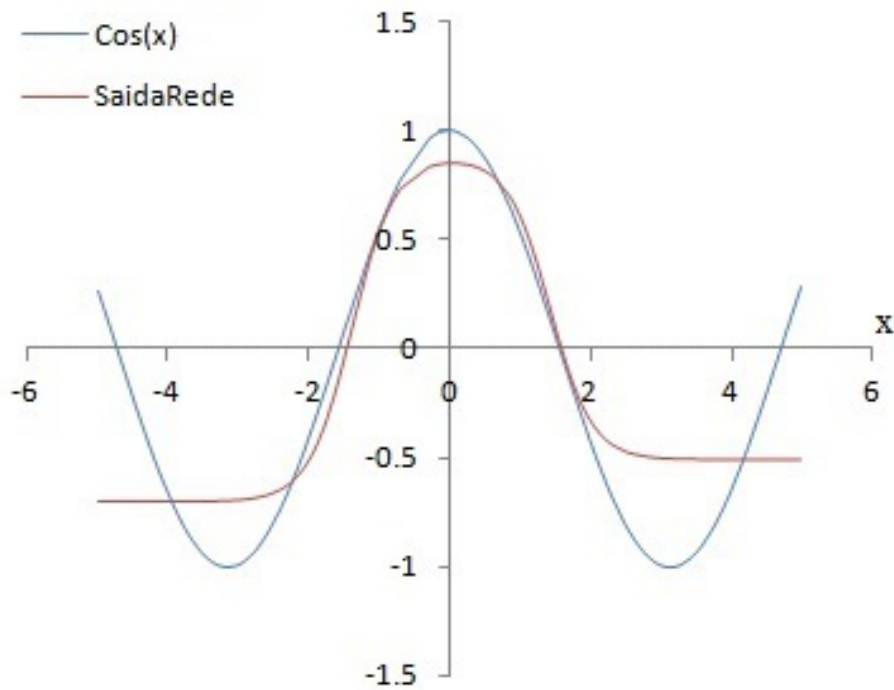


Fonte: O autor, 2016.

#### 3.4.1.2 Uma camada oculta com três neurônios.

Em função do mau desempenho da rede, o número de neurônios foi aumentado para três, e mantida uma camada oculta. Neste experimento, o intervalo considerado da variável  $x$  foi de  $\mathcal{I} = [-5.0, 5.0]$ . Houve uma pequena melhora (Figura 14), mas o erro da rede ainda permaneceu elevado, 0.330957 e o critério de parada atendido foi que o número máximo de épocas foi atingido.

Figura 14 - Cosseno e a saída da rede com topologia de uma unidade na camada de entrada, 3 neurônios na camada oculta e um neurônio na camada de saída.

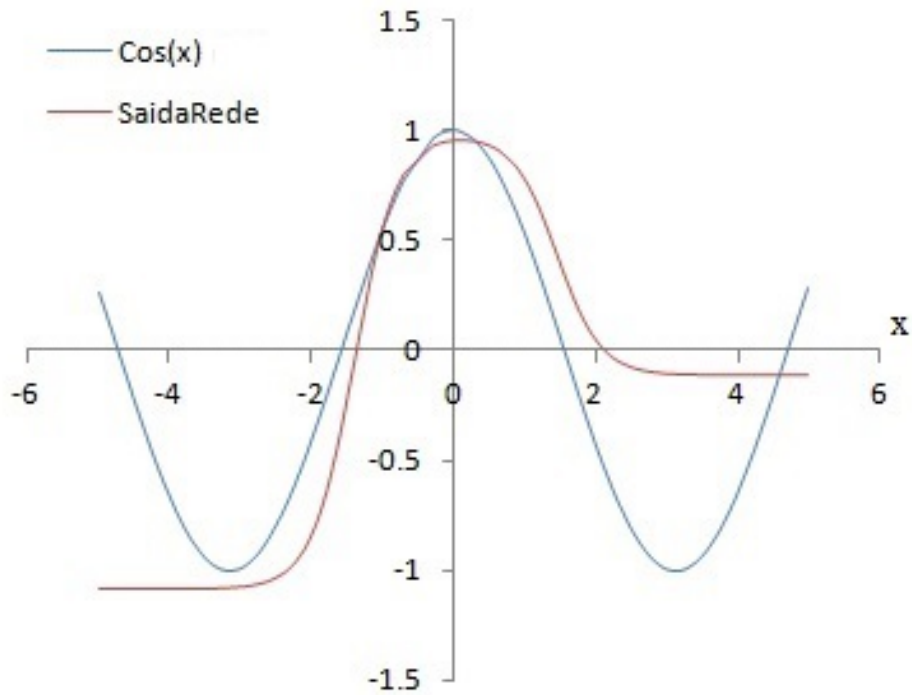


Fonte: O autor, 2016.

#### 3.4.1.3 Uma camada oculta com quatro neurônios.

Novamente o número de neurônios foi aumentado, agora para quatro na camada oculta. Neste experimento, o intervalo considerado da variável  $x$  foi de  $\mathcal{I} = [-5.0, 5.0]$ . O erro da rede permaneceu elevado, 0.323053 e o critério de parada foi que o número máximo de épocas foi atingido. A rede neuronal não conseguiu uma boa aproximação, como pode ser visto na Figura 15.

Figura 15 - Cosseno e a saída da rede com topologia de uma unidade na camada de entrada, 4 neurônios na camada oculta e um neurônio na camada de saída.

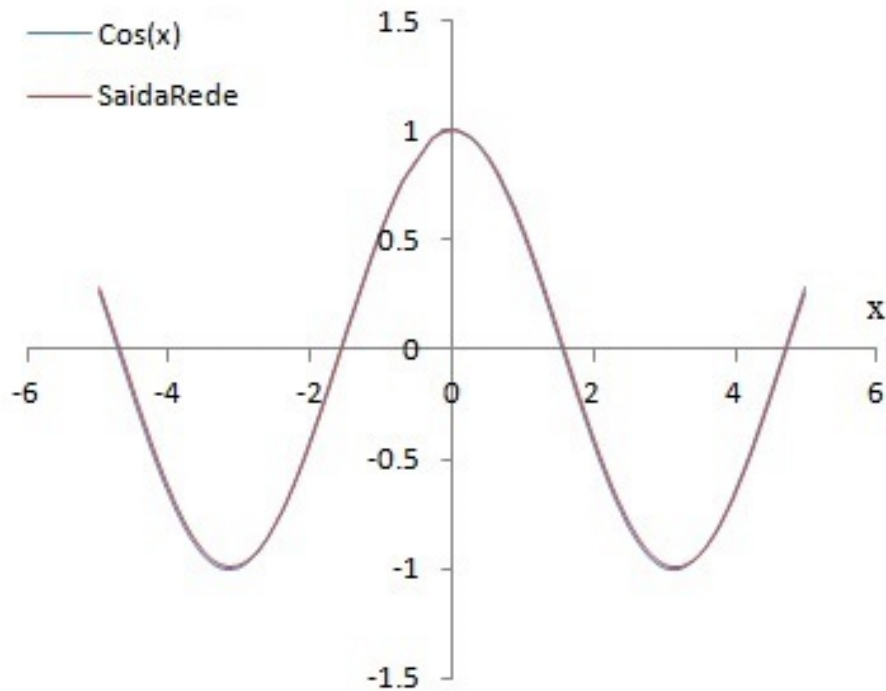


Fonte: O autor, 2016.

#### 3.4.1.4 Uma camada oculta com cinco neurônios.

Aumentado o número de neurônios para cinco na camada oculta e mantidos os mesmos valores para os outros parâmetros do experimento. Neste experimento, o intervalo considerado da variável  $x$  foi de  $\mathcal{I} = [-5.0, 5.0]$ . A rede convergiu, como pode ser visto na Figura 16, reduzindo o erro significativamente para 0.00295276. O critério de parada atendido foi que o erro quadrático médio foi menor que o erro mínimo.

Figura 16 - Cosseno e a saída da rede com topologia de uma unidade na camada de entrada, 5 neurônios na camada oculta e um neurônio na camada de saída.



Fonte: O autor, 2016.

Com a mesma topologia, uma unidade na camada de entrada, cinco neurônios na camada oculta e um neurônio na saída, o intervalo da variável  $x$  foi alterado para  $\mathcal{I} = [-8.5, 8.5]$ . Com esta alteração, o problema de saturação da função de ativação, observado no teste inicial, reapareceu, o que fez o algoritmo de treinamento levar os pesos a valores próximos de zero (Tabela 1) e a rede produzir saída constante e erro elevado, 0.676341. O critério de parada atendido foi que o número máximo de épocas foi atingido.

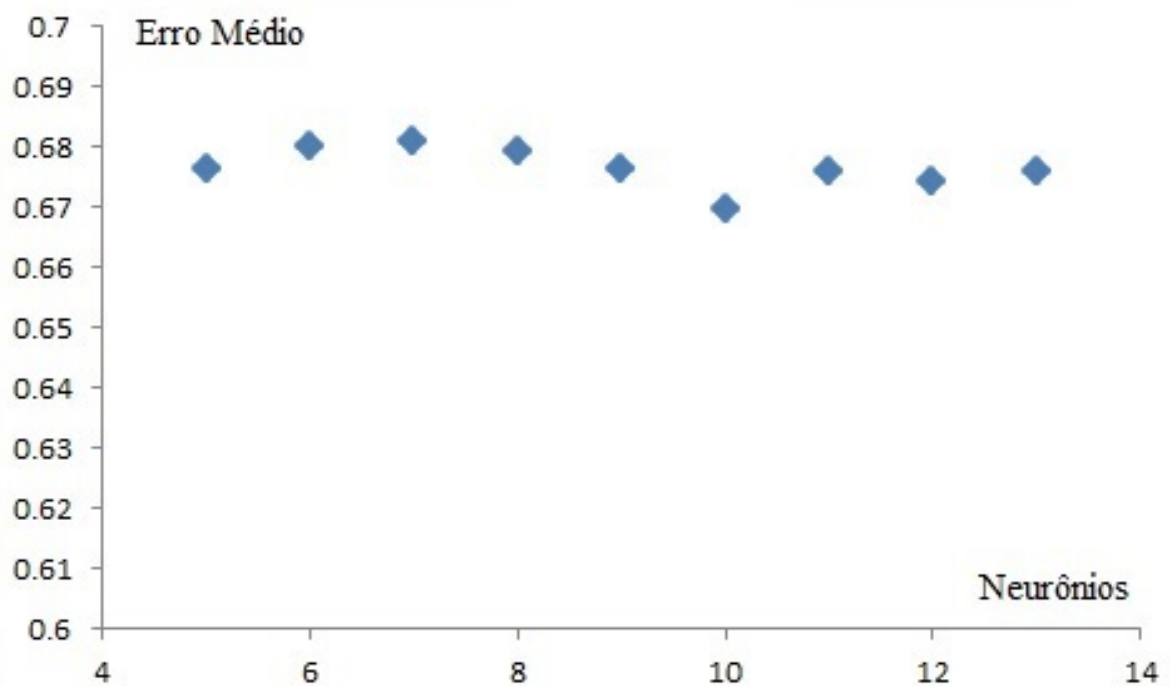
Mantidos o intervalo da variável  $x$  em  $\mathcal{I} = [-8.5, 8.5]$  e o mesmo conjunto de dados de entrada  $x$ , analisamos o comportamento do erro, frente a diferentes topologias da rede. Na Figura 17, apresentamos a evolução do erro quadrático médio em função do número de neurônios na camada oculta.

Tabela 1 - Pesos finais.

Pesos Finais
9.60532e-018
1.41964e-018
-1.02275e-017
-2.47785e-018
-3.42008e-017
1.74094e-017
2.81559e-017
-1.58744e-017
1.13132e-017
5.65322e-018
2.49798e-018
5.53889e-019
-2.81887e-017
3.54961e-018
5.58955e-018
-0.00265583

Fonte: O autor, 2016.

Figura 17 - Evolução do erro em função do aumento de neurônios.



Fonte: O autor, 2016.



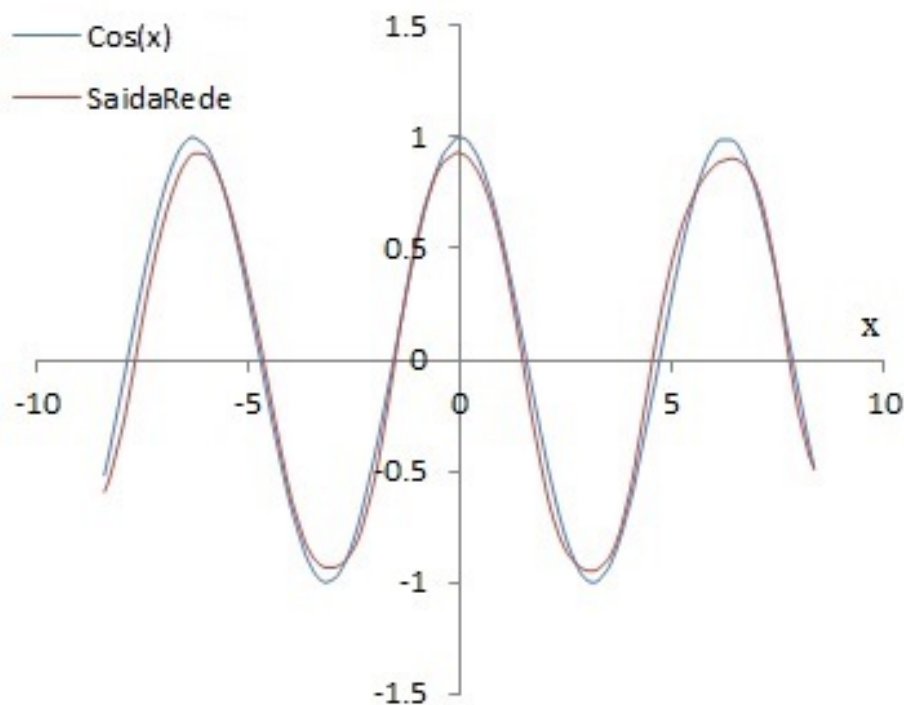
Observa-se que a estratégia de aumentar o número de neurônios não resolveu o problema da saturação, e o erro médio não foi reduzido, permanecendo entre 0.67 e 0.68. A estratégia seguinte foi aumentar o número de camadas.

### 3.4.2 Duas camadas ocultas.

#### 3.4.2.1 Duas camadas ocultas com cinco e dois neurônios.

Neste experimento, o intervalo considerado da variável  $x$  foi de  $\mathcal{I} = [-8.5, 8.5]$ . O erro médio foi de 0.0704072, com aproximação razoável da função cosseno (Figura 18). O critério de parada foi que o número máximo de épocas foi atingido.

Figura 18 - Cosseno e a saída da rede com topologia de uma unidade na camada de entrada, 5 neurônios na primeira camada oculta, 2 neurônios na segunda camada e um neurônio na camada de saída.



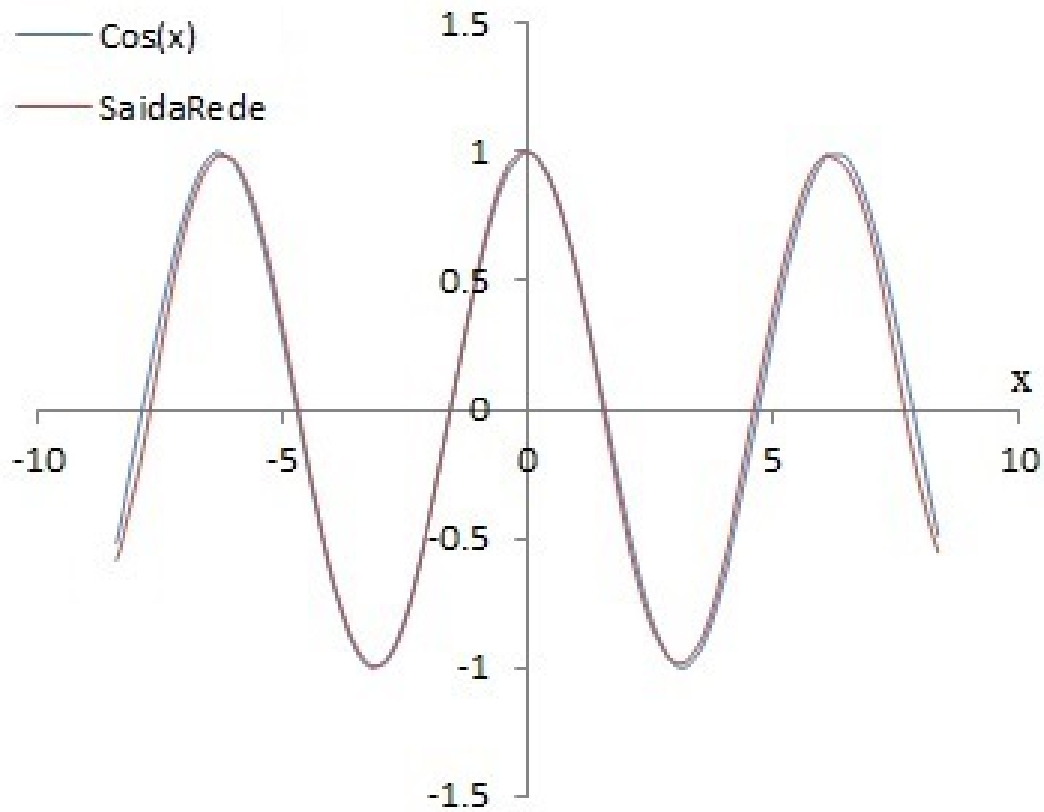
Fonte: O autor, 2016.

#### 3.4.2.2 Duas camadas ocultas com cinco e três neurônios.

Neste experimento, o intervalo considerado da variável  $x$  foi de  $\mathcal{I} = [-8.5, 8.5]$ . Com o aumento do número de neurônios da segunda camada de dois para três, o erro foi de 0.0243485, e a rede conseguiu uma aproximação um pouco melhor (Figura 19). O

critério de parada foi que o número máximo de épocas foi atingido.

Figura 19 - Cosseno e a saída da rede com topologia de uma unidade na camada de entrada, 5 neurônios na primeira camada oculta, 3 neurônios na segunda camada e um neurônio na camada de saída

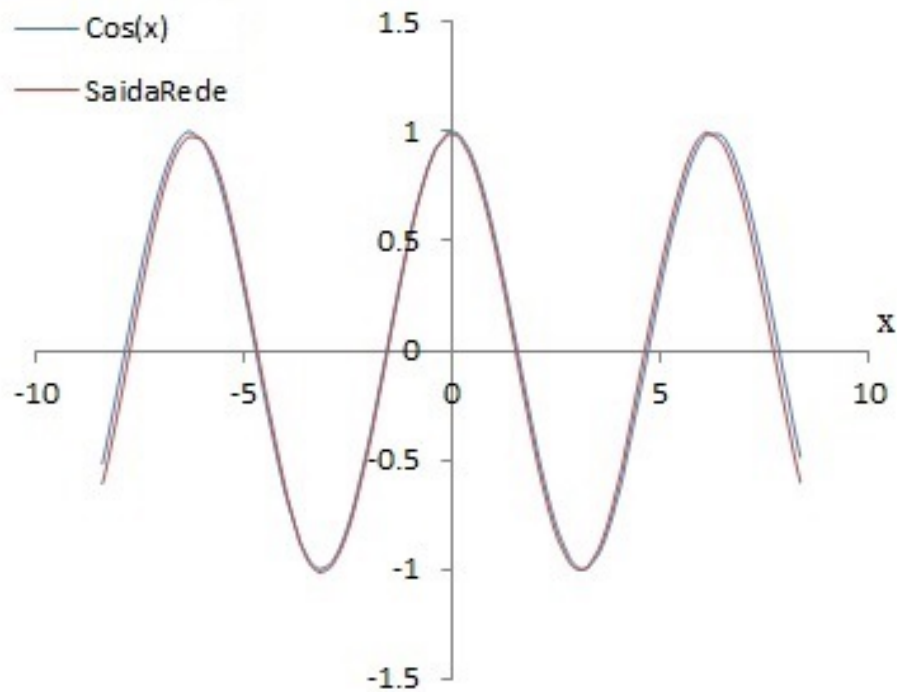


Fonte: O autor, 2016.

### 3.4.2.3 Duas camadas ocultas com cinco e quatro neurônios.

Neste experimento, o intervalo considerado da variável  $x$  foi de  $\mathcal{I} = [-8.5, 8.5]$ . A rede convergiu, como pode ser visto Figura 20, com erro de 0.00698692, que é menor que o erro mínimo estipulado.

Figura 20 - Cosseno e a saída da rede com topologia de uma unidade na camada de entrada, 5 neurônios na primeira camada oculta, 4 neurônios na segunda camada e um neurônio na camada de saída



Fonte: O autor, 2016.

#### 3.4.2.4 Duas camadas ocultas com sete e cinco neurônios.

Neste experimento, o intervalo considerado da variável  $x$  foi alterado para o intervalo não simétrico de  $\mathcal{I} = [-9.4, 10.7]$  e foi mantida a mesma topologia, uma unidade na camada de entrada, cinco neurônios na primeira camada oculta, quatro neurônios na segunda camada oculta e um neurônio na de saída. Nesse caso, não houve convergência.

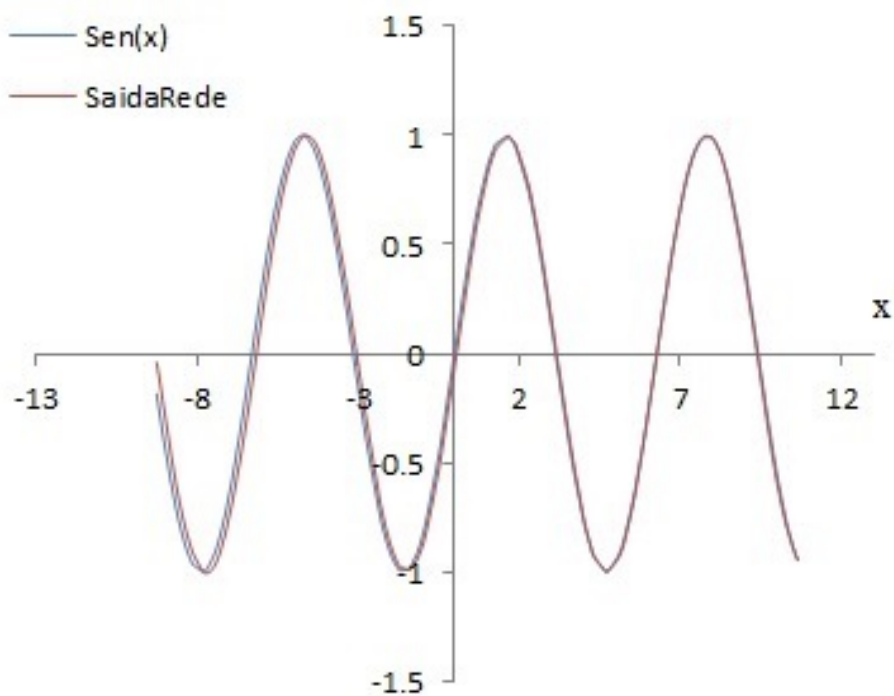
A topologia foi alterada conforme indicado na Tabela 2, e a rede convergiu, como pode ser visto na Figura 21, com sete neurônios na primeira camada oculta e cinco neurônios na segunda camada oculta.

Tabela 2 - Topologia de duas camadas e intervalo da entrada de  $[-9.4, 10.7]$ .

Topologia	Primeira Camada Oculta	Segunda Camada Oculta
T1	5	4
T2	6	4
T3	7	4
T4	7	5

Fonte: O autor, 2016.

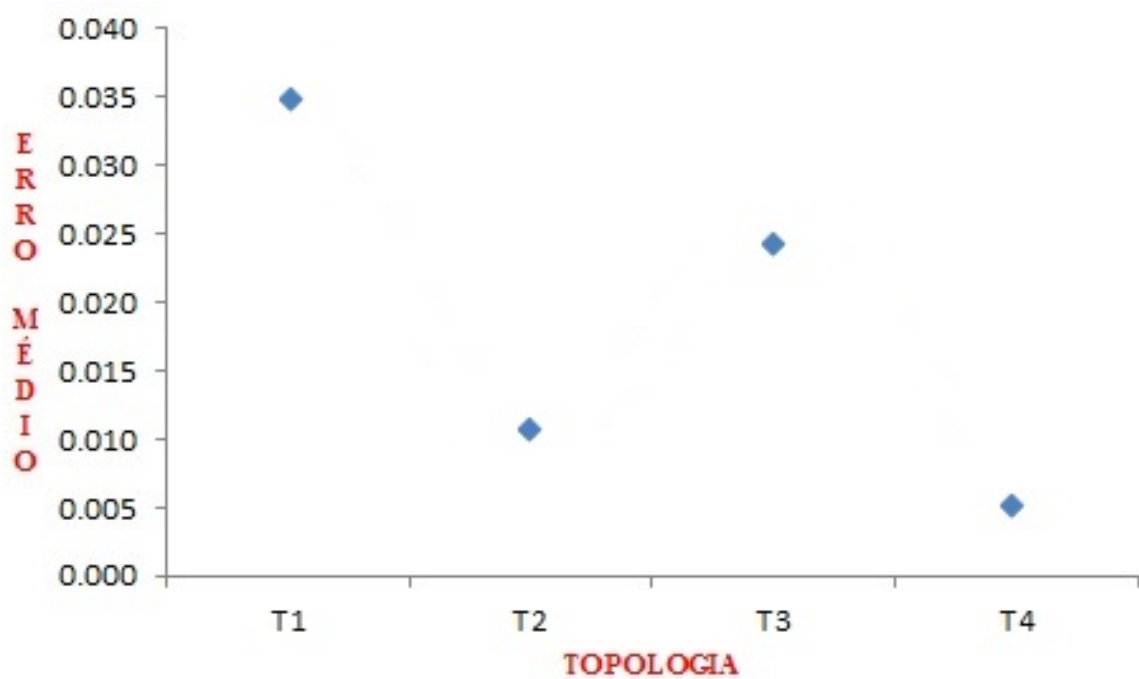
Figura 21 - Seno e a saída da rede com topologia de uma unidade na camada de entrada, 7 neurônios na primeira camada oculta, 5 neurônios na segunda camada e um neurônio na camada de saída



Fonte: O autor, 2016.

O erro quadrático médio foi de 0.006176 e sua evolução, em função das topologias descritas na Tabela 2, pode ser vista na Figura 22.

Figura 22 - Erro da rede em função da topologia com duas camadas ocultas e intervalo na entrada de  $[-9.4, 10.7]$



Fonte: O autor, 2016.

Alterado o intervalo da variável  $x$  para  $\mathcal{I} = [-15.708, 31.41]$ . Com a topologia de uma unidade na camada de entrada, sete neurônios na primeira camada oculta, cinco neurônios na segunda camada oculta e um neurônio na camada de saída, reaparecem os problemas de saturação e o da não convergência. A topologia foi então alterada conforme indicado na Tabela 3.

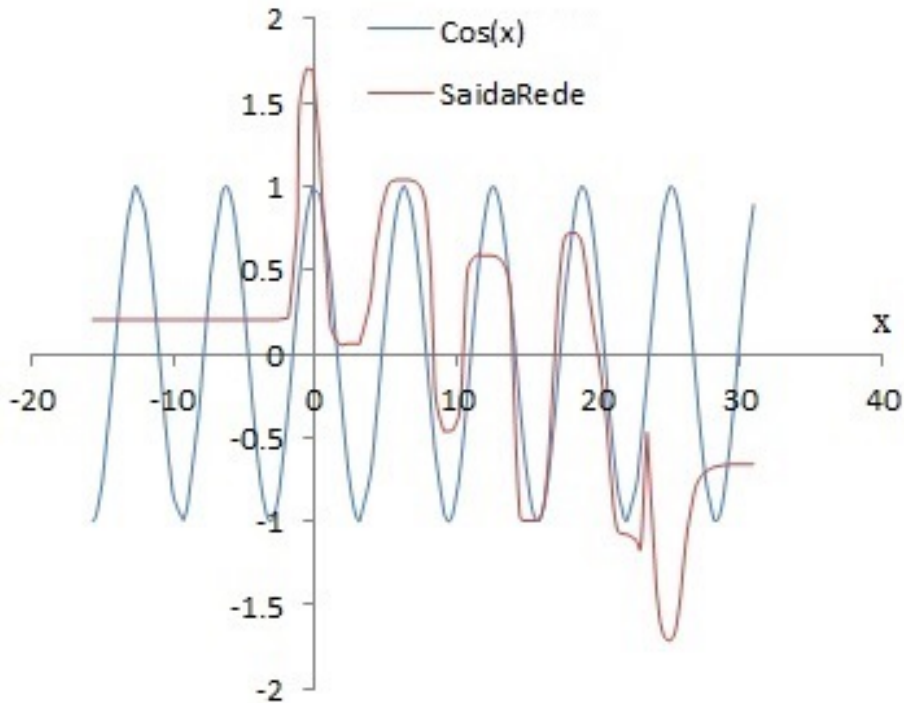
Tabela 3 - Topologia de duas camadas e intervalo de entrada de  $[-15.708, 31.41]$ .

Topologia	Primeira Camada Oculta	Segunda Camada Oculta
T1	7	5
T2	8	5
T3	8	6
T4	9	6
T5	9	7
T6	9	8
T7	10	7
T8	10	8
T9	10	9
T10	11	9
T11	11	10
T12	12	10
T13	13	10
T14	14	10

Fonte: O autor, 2016.

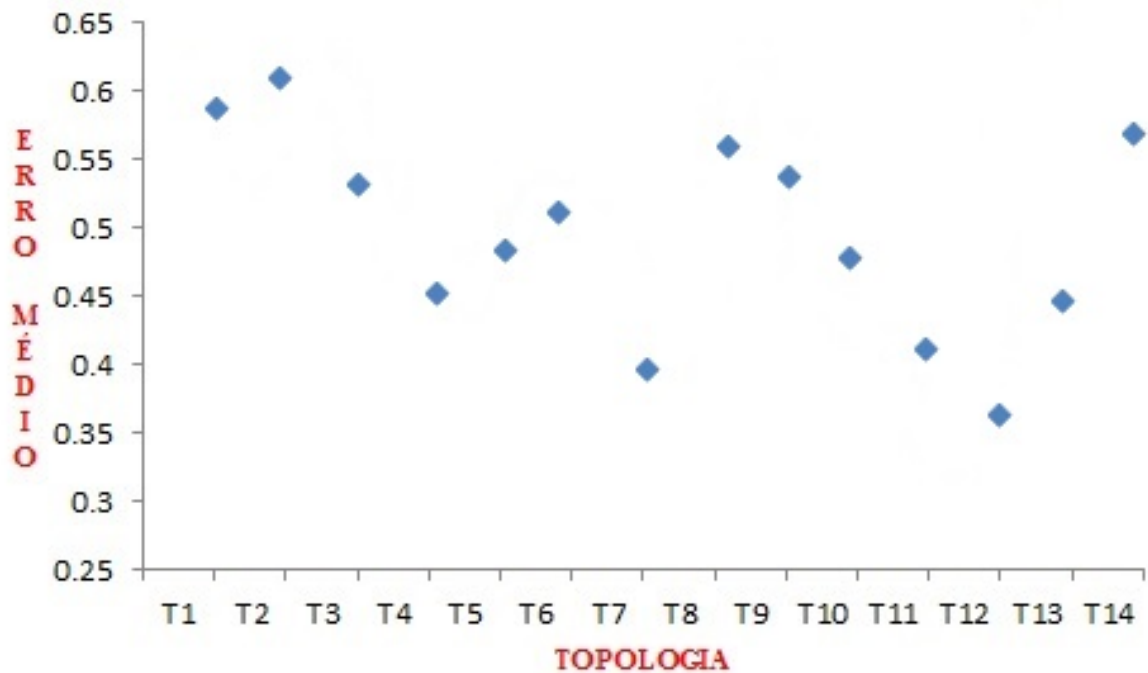
A rede não convergiu para nenhuma das topologias consideradas, parando sempre pelo número máximo de épocas e com erro superior ao mínimo estipulado. Na Figura 23, vemos o comportamento da rede para a topologia T14. A evolução do erro quadrático em função das topologias descritas na Tabela 3 pode ser vista na Figura 24.

Figura 23 - Cosseno e a saída da rede com topologia de uma unidade na camada de entrada, 14 neurônios na primeira camada oculta, 10 neurônios na segunda camada e um neurônio na camada de saída



Fonte: O autor, 2016.

Figura 24 - Erro da rede em função da topologia com duas camadas ocultas e intervalo na entrada de  $[-15.708, 31.41]$



Fonte: O autor, 2016.

### 3.4.3 Três camadas ocultas.

Neste experimento, o intervalo considerado da variável  $x$  foi mantido em  $\mathcal{I} = [-15.708, 31.41]$  e foram analisadas diferentes topologias. Foi acrescentada mais uma camada oculta e diferentes combinações do número de neurônios, em cada uma delas, foram testadas.

#### 3.4.3.1 Três camadas ocultas com quinze, dez e cinco neurônios.

Neste experimento, o intervalo considerado da variável  $x$  foi de  $\mathcal{I} = [-15.708, 31.41]$ . Acrescentar a terceira camada oculta não surtiu efeito, pois o erro médio permaneceu elevado. As diferentes topologias consideradas estão descritas na Tabela 4.

Tabela 4 - Topologia de três camadas.

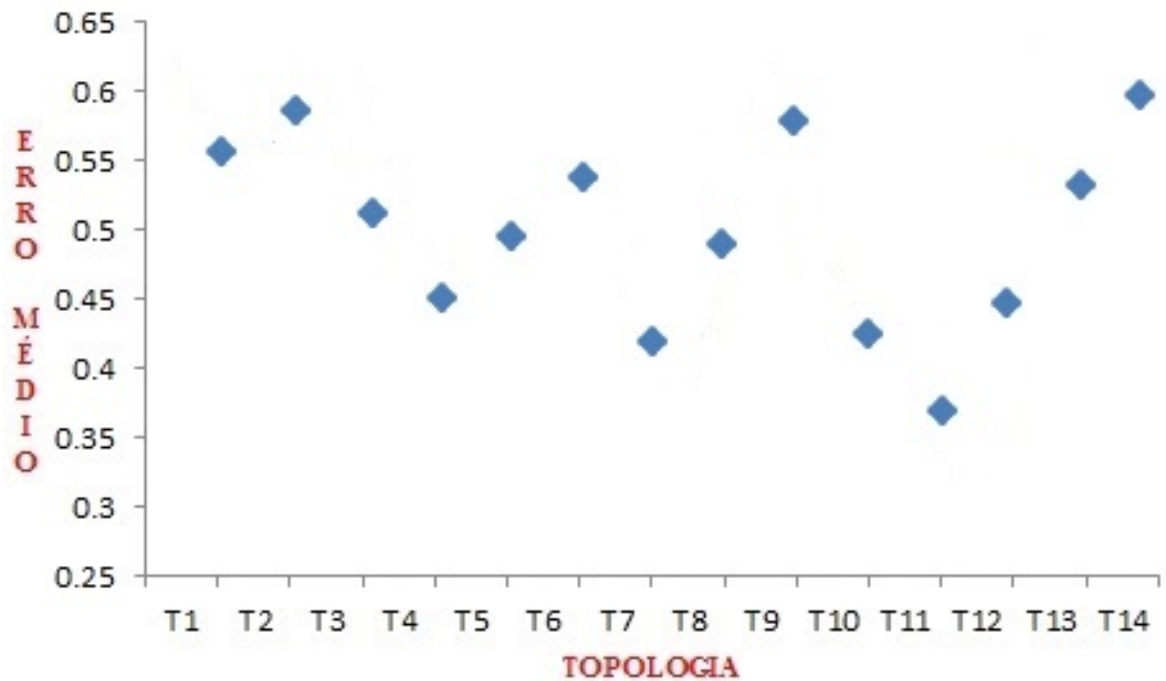
Topologia	Primeira Camada Oculta	Segunda Camada Oculta	Terceira Camada Oculta	Erro Médio
T1	14	10	2	0.620741
T2	14	10	3	0.568809
T3	14	10	4	0.561152
T4	14	10	5	0.62667
T5	15	10	5	0.612599

Fonte: O autor, 2016.

A evolução do erro quadrático, em função das topologias descritas na Tabela 4, pode ser vista na Figura 25.



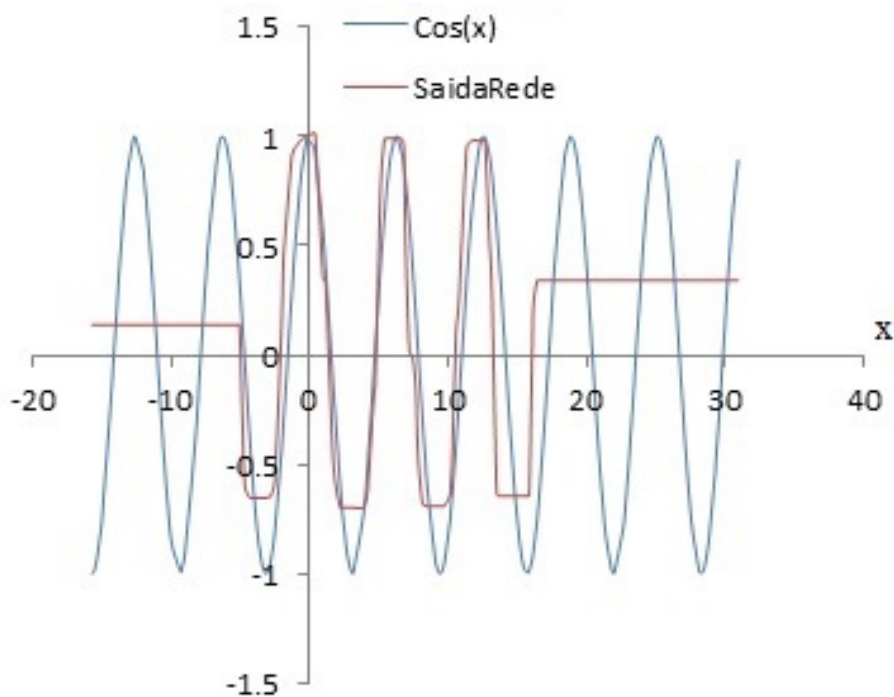
Figura 25 - Erro da rede em função da topologia com três camadas ocultas e intervalo na entrada de  $[-15.708, 31.41]$



Fonte: O autor, 2016.

Na Figura 26, vemos o comportamento da rede para a topologia T5.

Figura 26 - Cosseno e a saída da rede com topologia de uma unidade na camada de entrada, 15 neurônios na primeira camada oculta, 10 neurônios na segunda, 5 na terceira e um neurônio na camada de saída



Fonte: O autor, 2016.

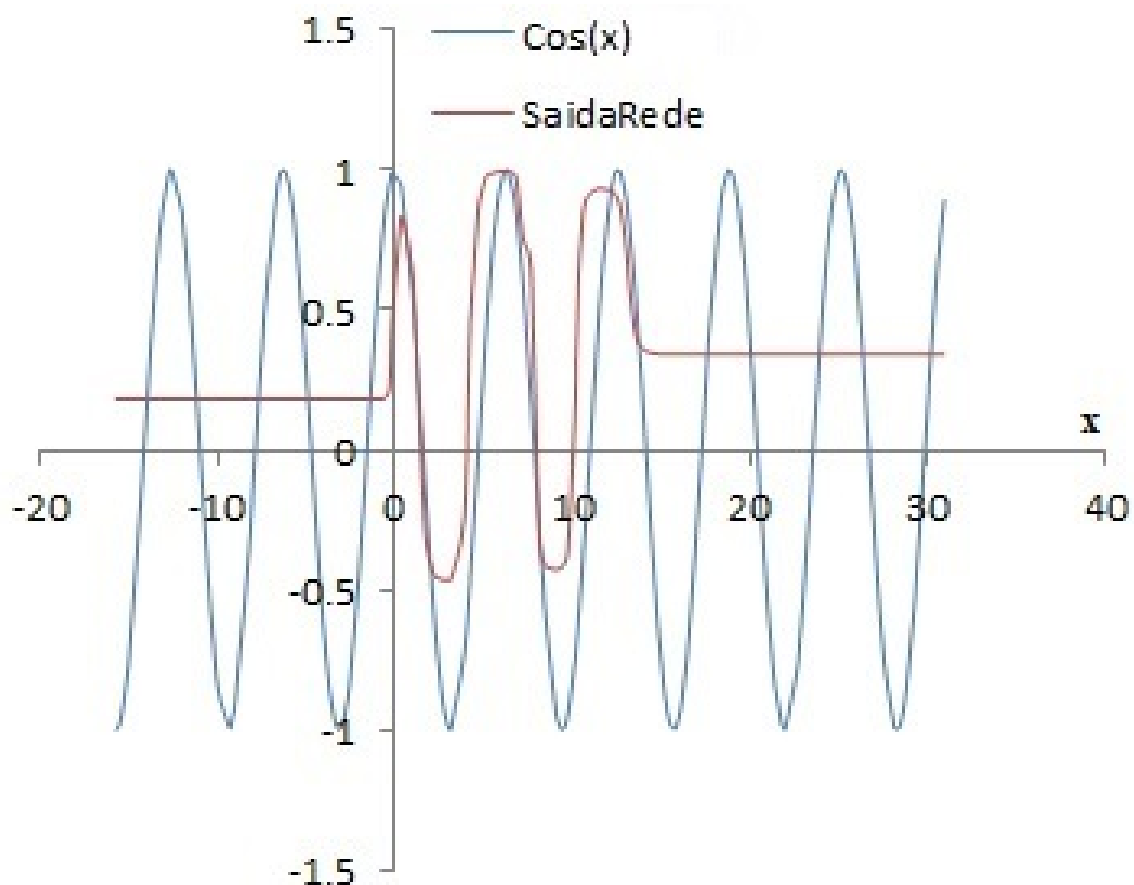
### 3.4.4 Quatro camadas ocultas.

O intervalo da variável  $x$   $\mathcal{I} = [-15.708, 31.41]$  foi mantido e foram analisadas diferentes topologias. Foi acrescentada mais uma camada oculta e diferentes combinações do número de neurônios, em cada uma delas, foram testadas.

#### 3.4.4.1 Quatro camadas ocultas com vinte, quinze, dez e cinco neurônios.

Neste experimento, o intervalo considerado da variável  $x$  foi de  $\mathcal{I} = [-15.708, 31.41]$ . Mais uma camada foi acrescentada, e usamos a topologia com uma unidade na camada de entrada, vinte neurônios na primeira camada oculta, quinze na segunda camada, dez na terceira, cinco na quarta e um neurônio na saída. Não houve redução do erro médio, 0.713708 e a rede, novamente, não convergiu, ver Figura 27. O critério de parada foi o de atingir número máximo de épocas.

Figura 27 - Cosseno e a saída da rede com topologia de uma unidade na camada de entrada, 20 neurônios na primeira camada oculta, 15 neurônios na segunda camada, 10 na terceira, 5 na quarta e um neurônio na camada de saída



Os experimentos com uma função conhecida permitiram que identifiquemos certas dificuldades a serem enfrentadas quando usamos redes neurais:

- Sensibilidade em relação aos valores de entrada da rede, que podem acarretar saturação da função de ativação. Por exemplo, a imagem da função tangente hiperbólica para o valor de 12.5 é 0.999999999972, que devido à representação numérica no computador torna-se 1, como já foi visto. Portanto, todos os valores superiores a 12.5 produzem na saída do neurônio, que tenha como função de ativação tangente hiperbólica, saída constante igual a 1.
- Dependência não linear do comportamento do erro em função do número de camadas. Somente a camada de saída tem o valor preciso do erro cometido. A última camada escondida recebe uma aproximação do erro, a penúltima uma aproximação da aproximação do erro, e assim por diante. Por isso o aumento no número de camadas não necessariamente reduz o erro da rede.
- Número elevado de neurônios das camadas ocultas pode fazer a rede decorar os dados de treinamento e ter baixo desempenho na fase de testes, não generalizando. Por outro lado, um número reduzido de neurônios pode acarretar em tempo excessivamente alto para rede encontrar uma solução ótima.
- A taxa de aprendizagem é outro parâmetro importantíssimo e tem grande influência no aprendizado da rede. Valores elevados podem provocar oscilações em torno de um valor ótimo. Valores reduzidos tornam o aprendizado lento. Seu valor normalmente varia entre 0.1 e 1. Nestes experimentos a taxa escolhida foi 0.1, que se mostrou adequada, não causando lentidão no treinamento, em função das características do problema, uma unidade na camada de entrada e um neurônio na camada de saída.

#### 4 A COLUNA DEPROPANIZADORA

Para inferir as duas propriedades, frações pesadas na produção de topo e frações leves na produção de fundo da torre depropanizadora, foram escolhidas dez variáveis de entrada para a rede: temperatura de topo, temperatura do prato 25, temperatura da vazão de topo, temperatura do prato 21, temperatura do prato 27, temperatura de fundo, vazão de vapor, pressão de topo, pressão de fundo e razão do refluxo de topo pela produção de propano.

A escolha destas variáveis, como estratégia inicial, levou em consideração que elas são utilizadas pela operação da unidade de craqueamento para, de forma indireta, inferir os parâmetros de qualidade da coluna depropanizadora. Pode-se dizer, que estes operadores possuem um “modelo” mental das colunas de destilação que eles operam, utilizando como entradas, para esta “modelagem”, as variáveis de processo da coluna (vazões, pressões e temperaturas). Como veremos, ao longo da descrição dos experimentos, algumas destas variáveis foram retiradas sem perda da qualidade das informações de entrada da rede neuronal.

A notação utilizada para as variáveis escolhidas, bem como a unidade e o intervalo de cada uma delas, estão indicados na Tabela 5.

Tabela 5 - Variáveis de entrada.

Sigla	Descrição	Intervalo	Unidade
T_T	Temperatura de Topo da Coluna	16.45 – 60.18	°C
T_P_25	Temperatura do Prato 25	21.74 – 93.19	°C
T_V_T	Temperatura da Vazão de Topo	15.05 – 46.47	°C
T_P_21	Temperatura do Prato 21	19.36 – 88.85	°C
T_P_27	Temperatura do Prato 27	21.66 – 95.05	°C
T_F	Temperatura de Fundo	15.89 – 98.43	°C
V_S	Vazão de Vapor	0.440 – 18.97	t/h
P_T	Pressão de Topo	0.170 – 18.97	kgf/cm <sup>2</sup>
P_F	Pressão de Fundo	2.400 – 18.58	kgf/cm <sup>2</sup>
R_R_P	Razão Refluxo de Topo e Produção de Propano	0.000 – 18.79	–
%PesadoProp	Percentual de Frações Pesadas no Propano	0.000 – 45.54	%
%LeveButano	Percentual de Frações Leves no Butano	0.010 – 33.52	%

Fonte: O autor, 2016.

#### 4.1 Tratamento inicial dos dados.

Os dados para os experimentos foram coletados da base de dados de processo da Refinaria Duque de Caxias. Inicialmente foram coletados os dados de laboratório, que serviram como critério para a coleta dos dados de processo. Os dados de laboratório representam os valores desejados das saídas da rede neuronal, e, em função da rotina de análises do laboratório da refinaria, eles são mais esparsos que os dados de processo. Com a data e a hora de cada dado de laboratório foram obtidos os dados de cada variável de processo, variáveis de entrada. Inicialmente, como não foi feito nenhum tratamento dos dados, os intervalos de variação de cada variável foram definidos a partir dos dados de laboratório e da consulta a especialistas.

Pelo intervalo de variação de cada variável foi possível perceber a existência de dados espúrios (outliers) na base de dados. Estes dados espúrios são oriundos de problemas de medição ou foram coletados em períodos em que a torre depropanizadora estava fora de sua condição normal de operação. São dados que não fornecerão informações consistentes para o treinamento da rede, e poderão até atrapalhar a sua convergência.

Por exemplo, o valor da pressão de topo (P\_T), na condição normal de operação, é de aproximadamente  $16 \text{ kgf/cm}^2$ , sendo aceitáveis valores de 14 a  $18 \text{ kgf/cm}^2$ . A temperatura de topo (T\_T) situa-se entre 50 e  $60 \text{ }^\circ\text{C}$ , e por isso valores abaixo de  $40 \text{ }^\circ\text{C}$  foram retirados. O mesmo tratamento foi dado para as demais variáveis de entrada.

Além disso, segundo um especialista da unidade de craqueamento, a torre não opera bem quando a razão entre refluxo de topo e produção de propano (R\_R\_P) está menor que 0.5. Então, todos os registros fora do intervalo de valores aceitáveis para a variável R\_R\_P foram retirados. Como consequência o intervalo das saídas, também, foi reduzido. Considerando os dados disponíveis após o tratamento, os intervalos das variáveis de entrada da rede e das saídas desejadas estão descritos na Tabela 6.

Tabela 6 - Variáveis de entrada após tratamento.

Sigla	Descrição	Intervalo	Unidade
T_T	Temperatura de Topo da Coluna	38.80 – 48.96	°C
T_P_25	Temperatura do Prato 25	56.57 – 91.40	°C
T_V_T	Temperatura da Vazão de Topo	33.85 – 42.54	°C
T_P_21	Temperatura do Prato 21	53.81 – 88.30	°C
T_P_27	Temperatura do Prato 27	57.41 – 93.98	°C
T_F	Temperatura de Fundo	68.95 – 98.43	°C
V_S	Vazão de Vapor	1.15 – 18.97	t/h
P_T	Pressão de Topo	13.90 – 18.00	kgf/cm <sup>2</sup>
P_F	Pressão de Fundo	10.46 – 18.58	kgf/cm <sup>2</sup>
R.R.P	Razão Refluxo de Topo e Produção de Propano	0.500 – 2.300	–
%PesadoProp	Percentual de Frações Pesadas no Propano	0.000 – 15.64	%
%LeveButano	Percentual de Frações Leves no Butano	0.010 – 19.65	%

Fonte: O autor, 2016.

## 4.2 Resultados

Não existe regra para a definição da topologia de uma rede neuronal. Também não existe especificação formal para o número de camadas ocultas e, tampouco, para o de neurônios de cada uma delas. As unidades de entrada e os neurônios da camada de saída são determinados em função do problema. O treinamento de uma rede neuronal deve começar com a topologia mais simples possível.

Existem algumas heurísticas muito utilizadas para determinar a quantidade de neurônios de uma rede com somente uma camada escondida, como por exemplo:

- Método de Kolmogorov

$$n_1 = 2 \cdot n + 1 \quad (9)$$

- Método de Fletcher-Gloss

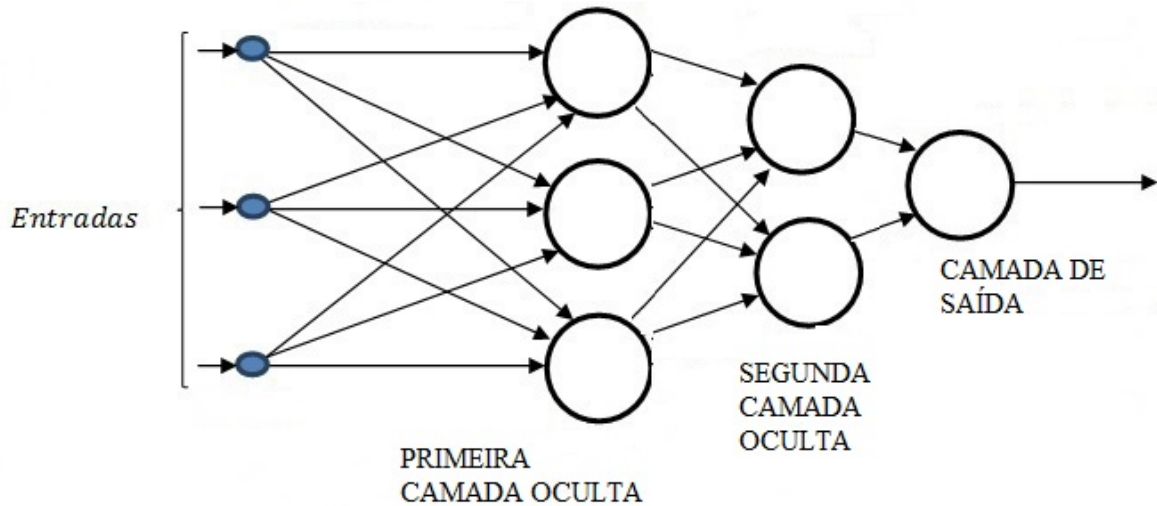
$$2 \cdot \sqrt{n} + n_2 \leq n_1 \leq 2 \cdot n + 1, \quad (10)$$

onde  $n$  é a quantidade de unidades de entrada da rede,  $n_1$  é a quantidade de neurônios da camada escondida e  $n_2$  é a de neurônios da camada de saída.

A sistemática utilizada, neste problema, foi a mesma dos experimentos das funções trigonométricas, cinco treinamentos em cada experimento e dez mil como número máximo de épocas. Em todos os experimentos em que as redes possuíam mais de uma camada oculta, o número de neurônio foi sempre decrescente a partir das unidades de en-

trada. Uma topologia de três unidades na camada de entrada, duas camadas ocultas com três e dois neurônios e um neurônio na camada de saída, esta representada na Figura 28.

Figura 28 - Exemplo de topologia dos experimentos



Fonte: O autor, 2016.

Agora a rede deverá identificar uma função de várias variáveis com valores vetoriais. Mais precisamente, inicialmente uma função de  $\mathbb{R}^{10}$  em  $\mathbb{R}^2$ . Durante os experimentos algumas variáveis de entrada foram retiradas e, além disso, a rede passou a identificar somente uma variável, isto é, um neurônio na camada de saída. O melhor resultado encontrado foi com cinco variáveis na entrada, isto é, uma função de  $\mathbb{R}^5$  em  $\mathbb{R}$ .

Os resultados serão apresentados divididos por topologia. Os erros quadráticos médios serão citados e o desempenho de cada topologia será apresentado de forma gráfica, sendo os valores desejados e as saídas calculadas, pela rede, no eixo  $y$ , e no eixo  $x$  a razão de refluxo de topo em relação à produção de propano (R\_R\_P).

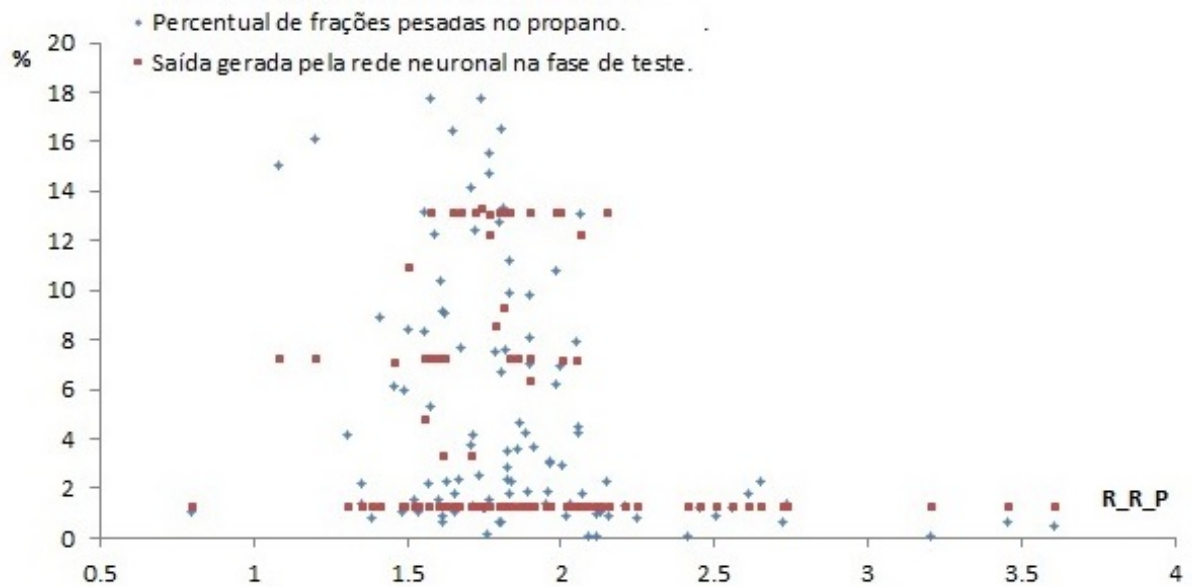
#### 4.3 Uma camada oculta e dois neurônios na camada de saída.

As heurísticas de topologia, dadas pelas Equações (9) e (10) não foram utilizadas “porque são recomendadas apenas para alguns tipos de problemas bem comportados. Porque desconsideram atributos relevantes da topologia como: a quantidade de dados, a complexidade do problema e a qualidade e disposição dos dados no seu espaço amostral” (FINOCCHIO, 2014). Com o objetivo de analisar a evolução do erro, de acordo com a Equação (5), começou-se com dez neurônios na camada escondida.

#### 4.3.1 Uma camada oculta com dez neurônios.

Este experimento utilizou uma topologia com dez unidades de entrada, uma camada oculta com dez neurônios e dois neurônios na camada de saída. A rede neuronal apresentou erro quadrático médio elevado, 3.66312 no conjunto da fase de treinamento e 3.85428 na de teste, e a rede não fez uma boa aproximação como pode ser constatado nas Figuras 29 e 30. O critério de parada atendido foi que o número máximo de épocas foi atingido.

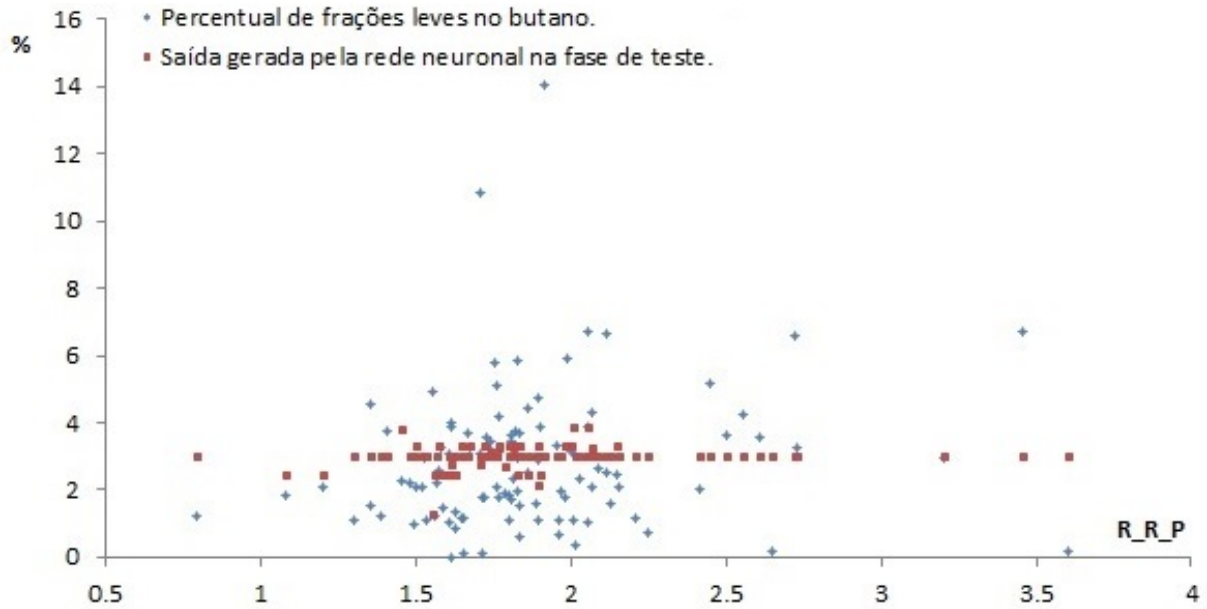
Figura 29 - Saída da rede neuronal e o valor esperado de percentual de frações pesadas no propano, em relação à razão de refluxo de topo e produção de propano (R.R.P em %), com topologia de 10 unidades na camada de entrada, 10 neurônios na camada oculta e 2 neurônios na camada de saída



Fonte: O autor, 2016.



Figura 30 - Saída da rede neuronal e o valor esperado de percentual de frações leves no butano, em relação à razão de refluxo de topo e produção de propano (R\_R\_P em %), com topologia de 10 unidades na camada de entrada, 10 neurônios na camada oculta e 2 neurônios na camada de saída

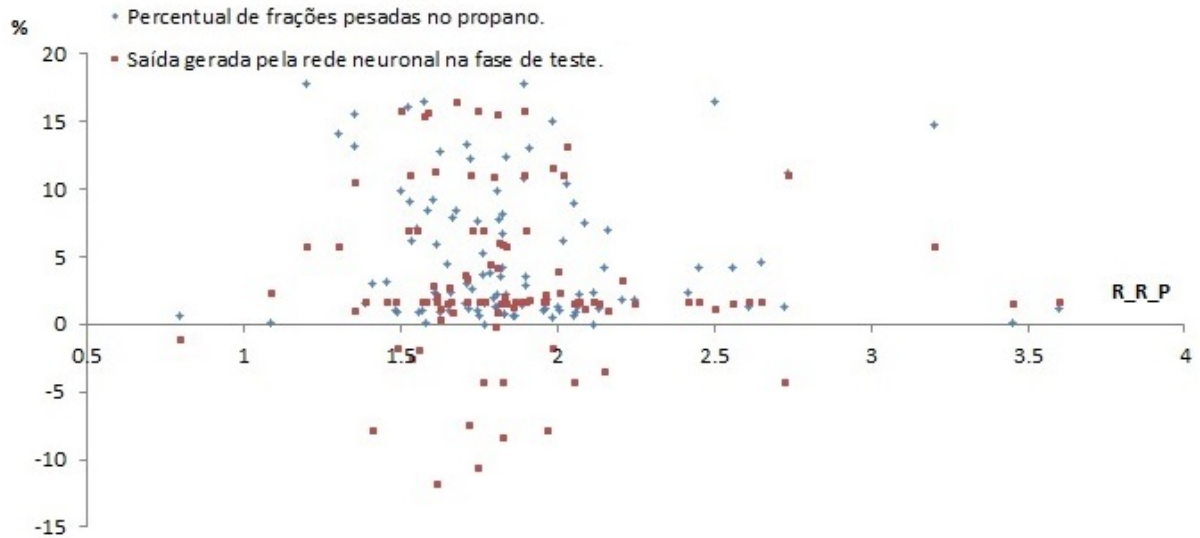


Fonte: O autor, 2016.

#### 4.3.2 Uma camada oculta com quinze neurônios.

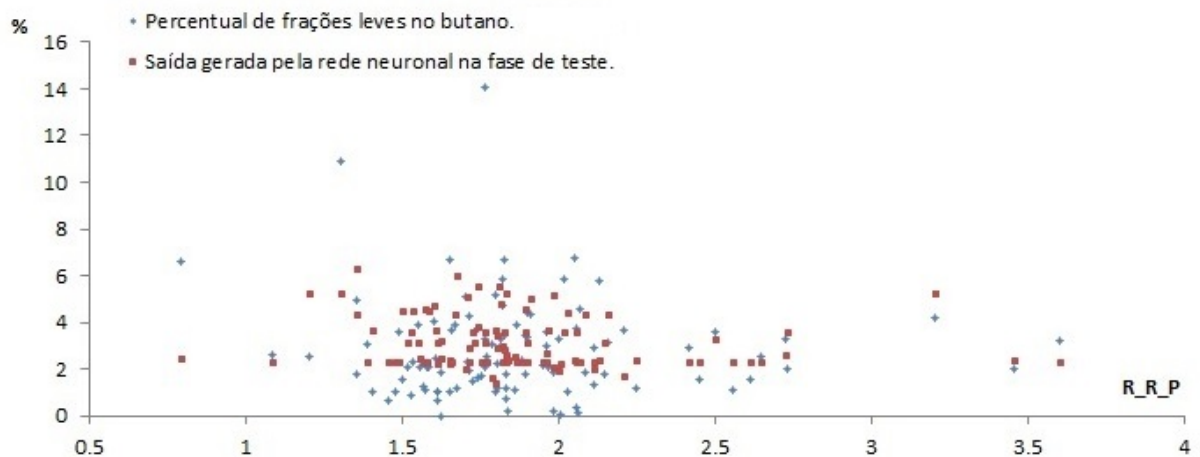
Como na estratégia utilizada no treinamento com as funções trigonométricas, foi aumentado o número de neurônios da camada oculta de dez para quinze neurônios. Este experimento utilizou uma topologia com dez unidades de entrada, uma camada oculta com quinze neurônios e dois neurônios na camada de saída. A rede neuronal novamente apresentou erro quadrático médio elevado, 3.8534 no conjunto da fase de treinamento e 6.16825 no conjunto da fase de teste, e a rede não fez uma boa aproximação como pode ser visto nas Figuras 31 e 32. O critério de parada atendido foi que o número máximo de épocas foi atingido.

Figura 31 - Saída da rede neuronal e o valor esperado de percentual de frações pesadas no propano, em relação à razão de refluxo de topo e produção de propano (R\_R\_P em %), com topologia de 10 unidades na camada de entrada, 15 neurônios na camada oculta e 2 neurônios na camada de saída.



Fonte: O autor, 2017.

Figura 32 - Saída da rede neuronal e o valor esperado de percentual de frações leves no butano, em relação à razão de refluxo de topo e produção de propano (R\_R\_P em %), com topologia de 10 unidades na camada de entrada, 15 neurônios na camada oculta e 2 neurônios na camada de saída.



Fonte: O autor, 2017.

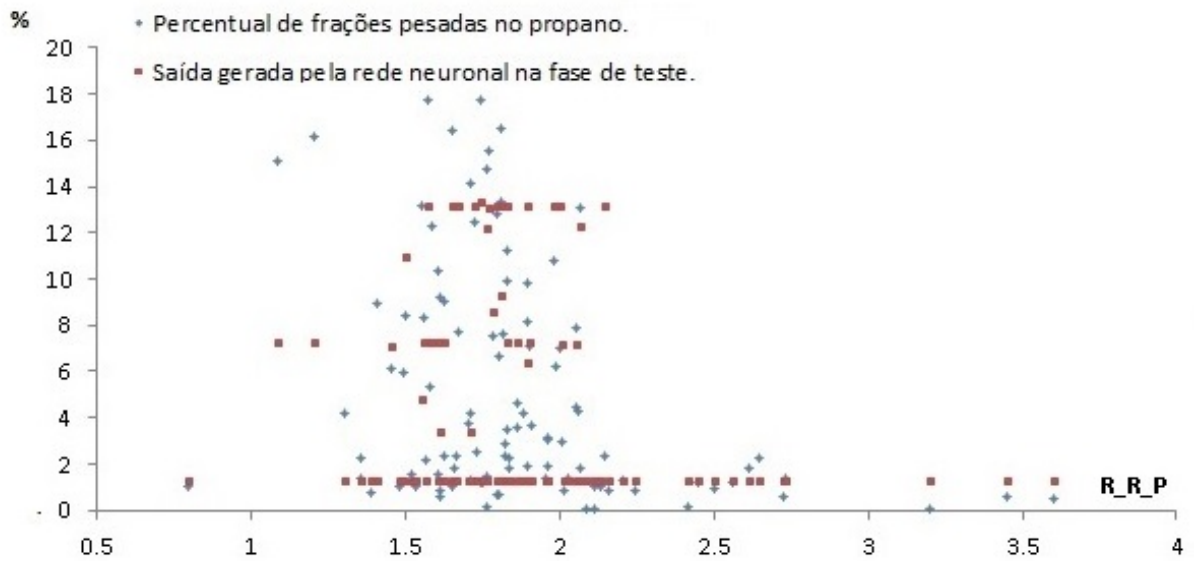
#### 4.3.3 Uma camada oculta com vinte neurônios.

Aumentado o número de neurônios da camada oculta de quinze para vinte neurônios. Este experimento utilizou uma topologia com dez unidades de entrada, uma

camada oculta com vinte neurônios e dois neurônios na camada de saída.

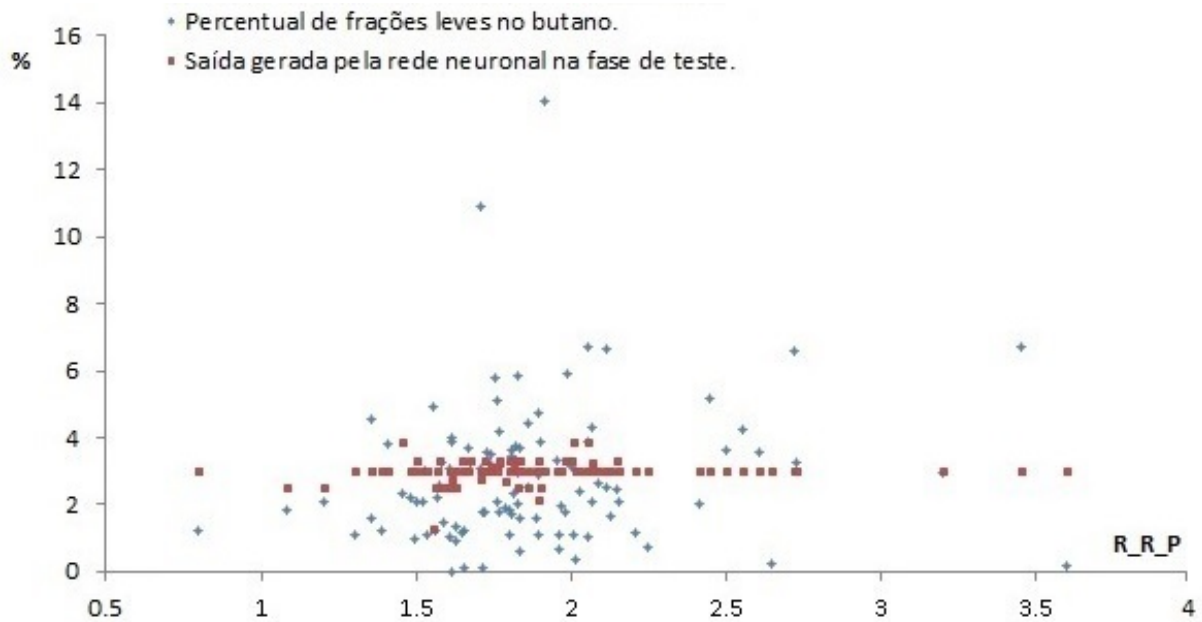
A rede neuronal não convergiu, ver Figuras 33 e 34, e o erro quadrático médio aumentou consideravelmente, foi de 73.7068 no conjunto da fase de treinamento e de 73.9816 no conjunto da fase de teste. O critério de parada atendido foi que o número máximo de épocas foi atingido.

Figura 33 - Saída da rede neuronal e o valor esperado de percentual de frações pesadas no propano, em relação à razão de refluxo de topo e produção de propano (R\_R\_P em %), com topologia de 10 unidades na camada de entrada, 20 neurônios na camada oculta e 2 neurônios na camada de saída.



Fonte: O autor, 2017.

Figura 34 - Saída da rede neuronal e o valor esperado de percentual de frações leves no butano, em relação à razão de refluxo de topo e produção de propano (R\_R\_P em %), com topologia de 10 unidades na camada de entrada, 20 neurônios na camada oculta e 2 neurônios na camada de saída.



Fonte: O autor, 2017.

Devido ao problema de separação imperfeita na torre depropanizadora, em função de ajustes inadequados das variáveis de processo, a produção de topo da coluna, propano, pode estar dentro da especificação, isto é, respeitando o limite máximo de 4% de frações pesadas e a produção de fundo, butano, não estar, portanto, contendo mais frações leves que o desejável. Desta forma, pode-se concluir que as saídas da rede, frações pesadas na produção de topo e frações leves na produção de fundo, não são correlacionadas, permitindo que elas sejam calculadas de forma independente, por redes distintas. Então, optou-se por calcular, inicialmente, as frações pesadas no produto de topo, por ser esta a especificação mais importante da torre depropanizadora.

Constatou-se que existiam poucas amostras de treinamento em que a variável de entrada R\_R\_P é menor que 1.4. Desta forma, o intervalo desta entrada ficou definido em [1.4, 2.3]. Valores muito acima do máximo tolerável (4%) de frações pesadas no propano indicam problema operacional na coluna depropanizadora, por isso, amostras de treinamento, que possuam estes valores pouco ou nada contribuem para um bom desempenho da rede neuronal. Então, decidiu-se por descartar amostras cujos valores de %PesadoProp sejam superiores a 6%, considerando que qualquer valor acima de 4% é considerado fora de especificação, independente do seu valor numérico.

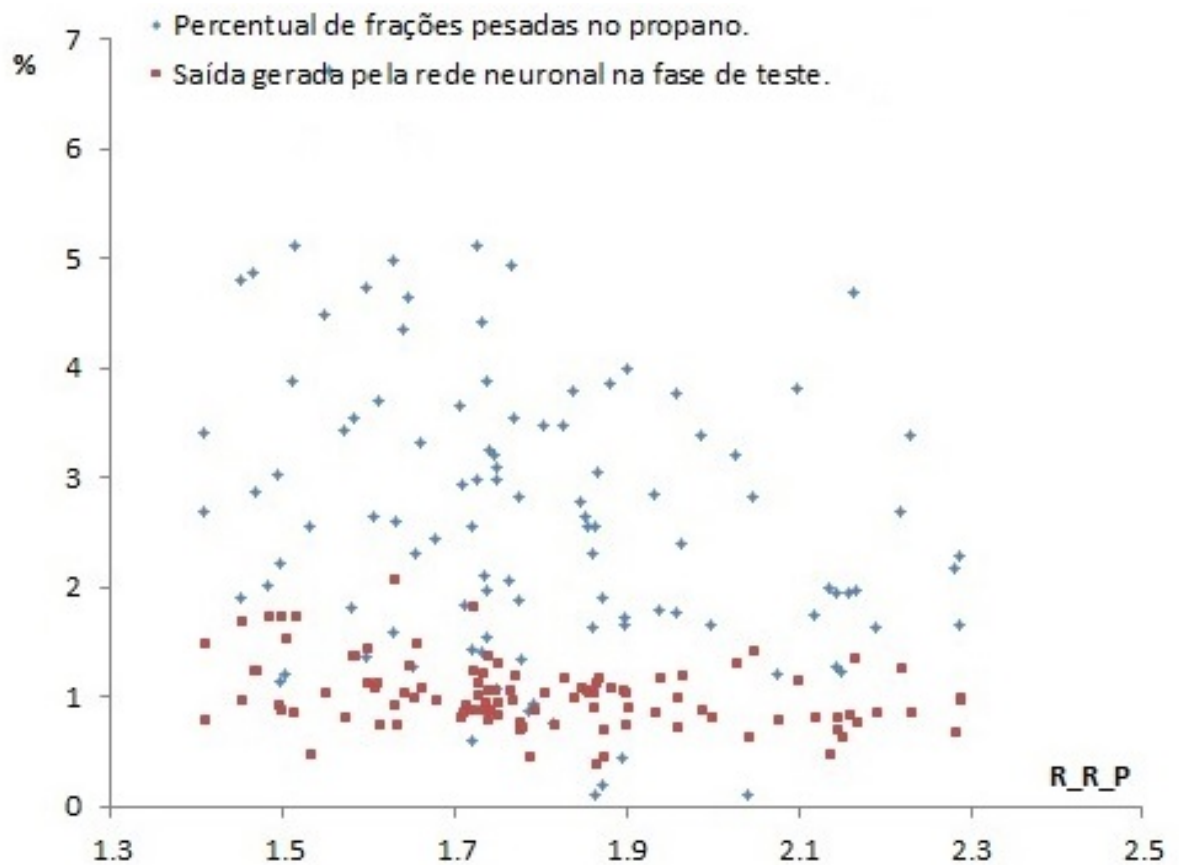
#### 4.4 Uma camada oculta e um neurônio na saída.

Este experimento utilizou uma topologia de dez unidades na camada de entrada, quinze neurônios na camada oculta e um neurônio na camada de saída, calculando somente as frações pesadas no propano.

##### 4.4.1 Uma camada oculta com quinze neurônios.

Houve redução do erro quadrático médio, sendo 1.31871 no conjunto da fase de treinamento e 1.95099 na de teste. Mas o desempenho da rede ainda ficou abaixo do desejado, como pode ser visto na Figura 35. O critério de parada atendido foi que o número máximo de épocas foi atingido.

Figura 35 - Saída da rede neuronal e o valor esperado de percentual de frações pesadas no propano, em relação à razão de refluxo de topo e produção de propano (R.R.P em %), com topologia de 10 unidades na camada de entrada, 15 neurônios na camada oculta e um neurônio na camada de saída.



Fonte: O autor, 2017.

#### 4.4.2 Uma camada oculta com trinta neurônios.

Aumentado o número de neurônios da camada oculta de quinze para trinta neurônios. Este experimento utilizou uma topologia com dez unidades na camada de entrada, trinta neurônios na oculta e um neurônio na de saída. Houve aumento do erro quadrático médio. O menor erro no conjunto da fase de treinamento foi de 2.22646, e no conjunto da fase de teste de 2.226464.

A rede não fez boa aproximação e gerou saídas com percentuais negativos. O critério de parada atendido foi que o número máximo de épocas foi atingido. Como o aumento do número de neurônios da camada oculta resultou em aumento do erro quadrático médio optou-se por aumentar de uma para duas camadas ocultas.

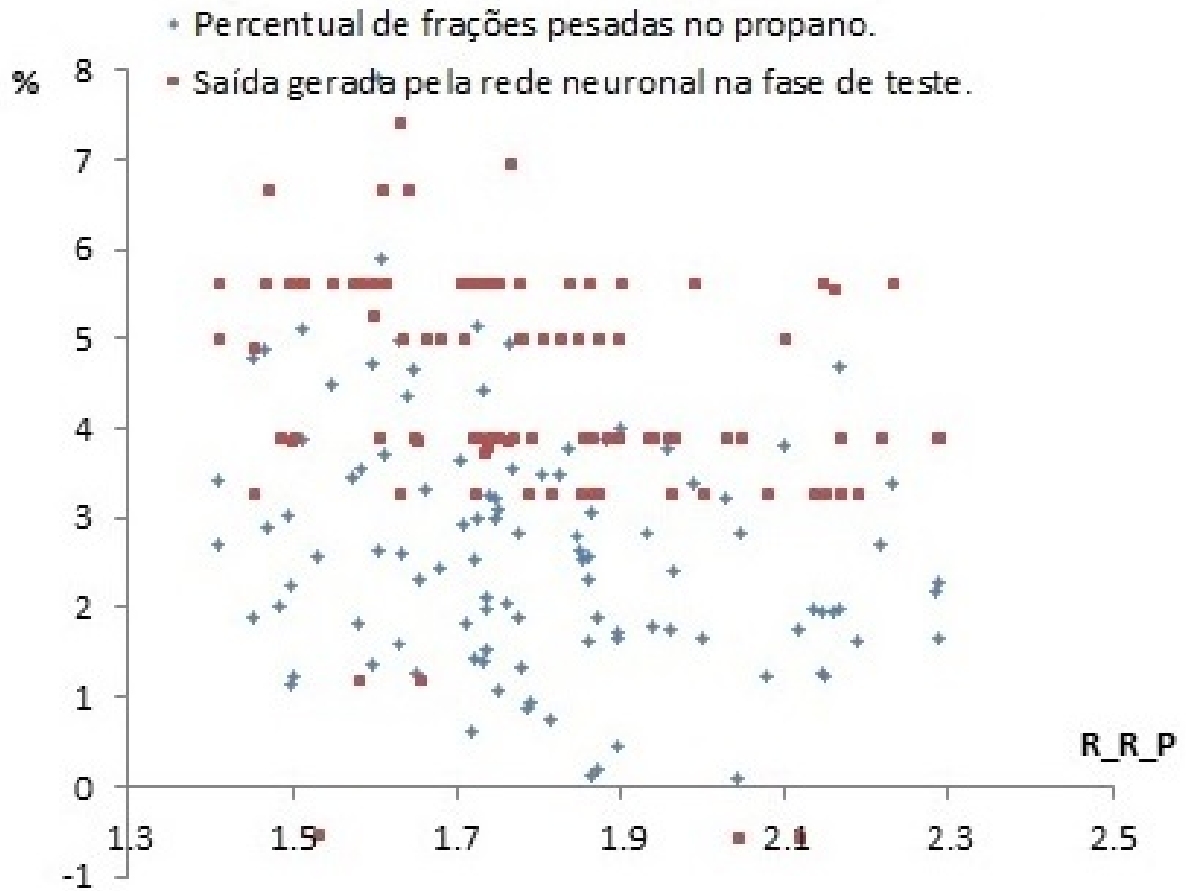
#### 4.5 **Duas camadas ocultas e um neurônio na saída.**

Este experimento utilizou uma topologia de dez unidades na camada de entrada, trinta neurônios na primeira camada oculta, quinze neurônios na segunda camada oculta e um neurônio na de saída.

##### 4.5.1 Duas camadas ocultas com trinta e quinze neurônios.

Houve pequena redução do erro quadrático médio. O menor erro no conjunto da fase de treinamento foi de 2.08196, e no conjunto da fase de teste de 2.23343. Apareceram, novamente, percentuais negativos nos valores calculados na camada de saída da rede (ver Figura 36). O critério de parada atendido foi que o número máximo de épocas foi atingido.

Figura 36 - Saída da rede neuronal e o valor esperado de percentual de frações pesadas no propano, em relação à razão de refluxo de topo e produção de propano (R.R.P em %), com topologia de 10 unidades na camada de entrada, 30 neurônios na primeira camada oculta, 15 neurônios na segunda camada e um neurônio na camada de saída.



Fonte: O autor, 2017.

#### 4.6 Três camadas ocultas e um neurônio na saída.

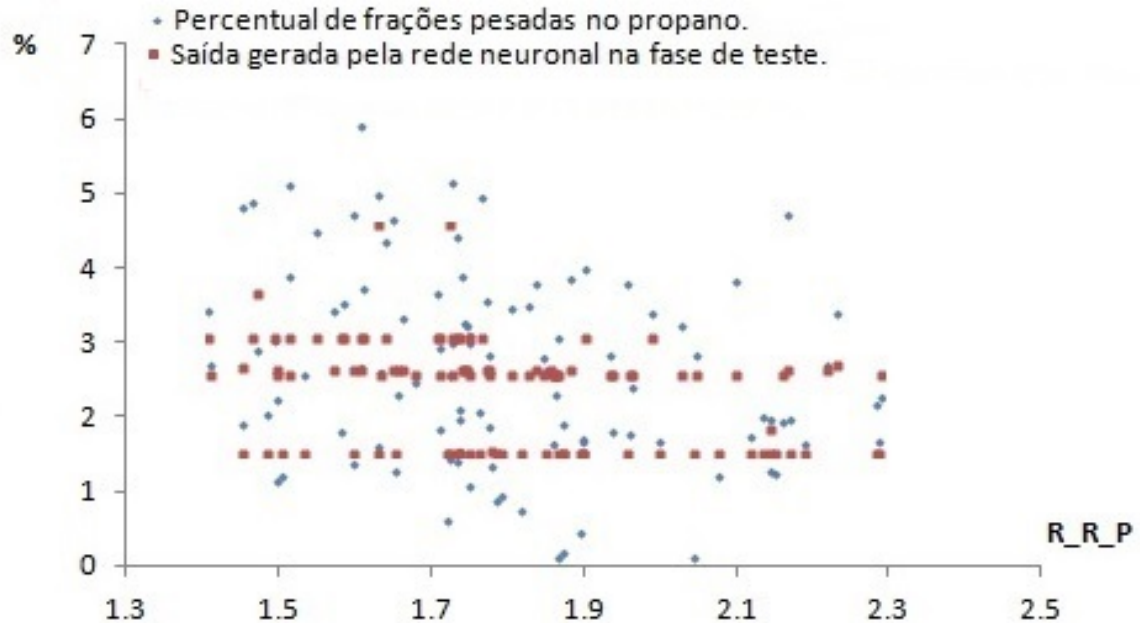
Aumentado o número de camadas ocultas de duas para três camadas.

##### 4.6.1 Três camadas ocultas com trinta, quinze e cinco neurônios.

Este experimento utilizou uma topologia de dez unidades na camada de entrada, trinta neurônios na primeira camada oculta, quinze neurônios na segunda, cinco na terceira e um neurônio na camada de saída. O menor erro no conjunto da fase de treinamento foi de 1.03191, e no conjunto da fase de teste de 0.952044. Com esta topologia a rede não calculou valores negativos, mas os valores ficaram concentrados em duas regiões,

1.5% e 3%, como pode ser visto na Figura 37.

Figura 37 - Saída da rede neuronal e o valor esperado de percentual de frações pesadas no propano, em relação à razão de refluxo de topo e produção de propano (R\_R\_P em %), com topologia de 10 unidades na camada de entrada, 30 neurônios na primeira camada oculta, 15 neurônios na segunda camada, 5 na terceira e um neurônio na camada de saída.



Fonte: O autor, 2017.

Desta forma, a rede neuronal não representou adequadamente os valores da propriedade que se deseja inferir, frações pesadas no propano. O critério de parada atendido foi que o número máximo de épocas foi atingido.

#### 4.7 Retirando variáveis de entrada da rede.

A decisão de calcular os parâmetros de qualidade, frações pesadas no propano e de leves no butano, em redes distintas, possibilitou a revisão das variáveis de entrada. Segundo especialistas da unidade de craqueamento, espera-se que as variáveis mais próximas do fundo da coluna depropanizadora tenham mais influência no parâmetro de qualidade frações leves no butano, e pouca influência na produção de topo. O mesmo pode-se dizer das variáveis mais próximas do topo em relação as frações pesadas no propano. Então foi verificado o desempenho da rede sem algumas variáveis, através da retirada destas e posterior avaliação do comportamento do erro quadrático médio. Como critério de retirada das variáveis, iniciou-se pelas variáveis de fundo, sendo a escolha entre estas em função da localização física dos pontos de medição ao longo da coluna, seguindo a sequência das variáveis do fundo para o topo da coluna.

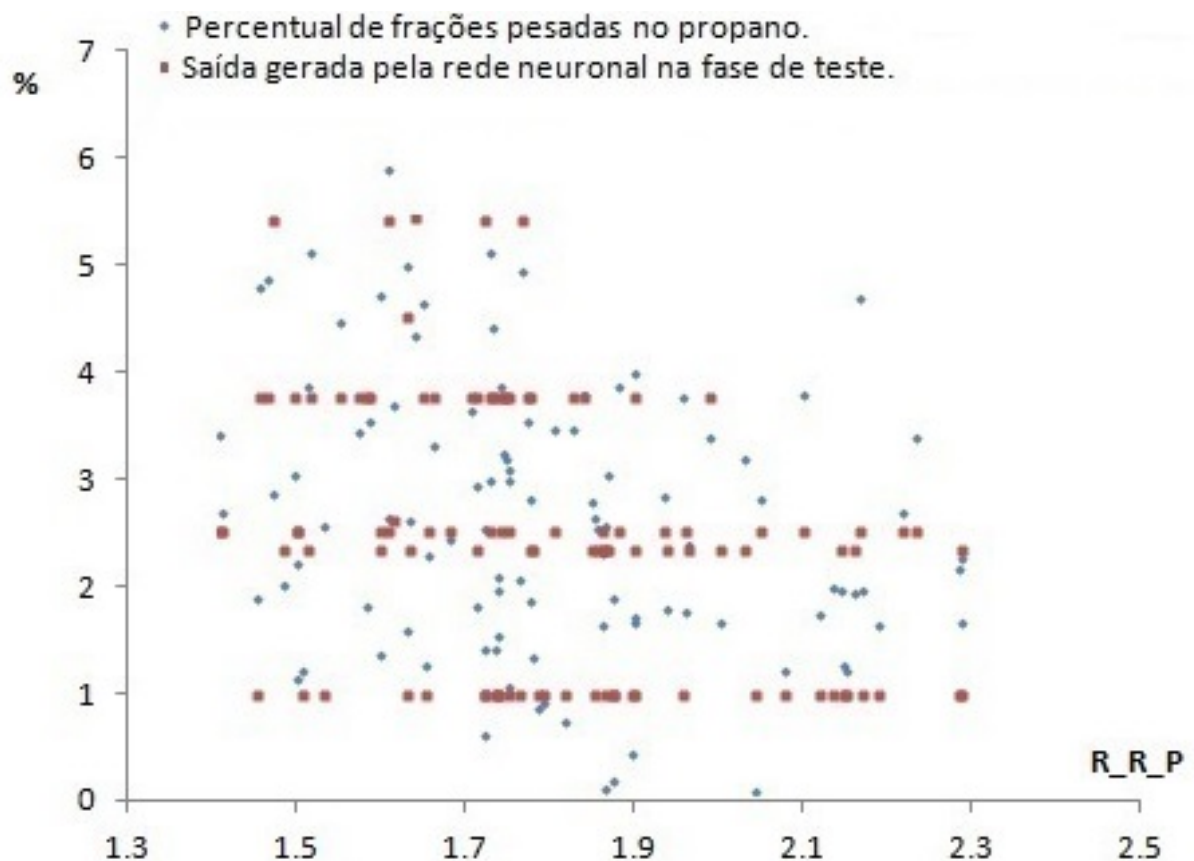


#### 4.7.1 Rede com nove unidades de entrada.

Foi retirada a variável de entrada pressão de fundo (P\_F) da coluna depropanizadora. Este experimento utilizou uma topologia de nove unidades na camada de entrada, trinta neurônios na primeira camada oculta, quinze neurônios na segunda, cinco na terceira e um na camada de saída. O menor erro no conjunto da fase de treinamento foi de 1.01584, e no conjunto da fase de teste de 0.967757.

O erro foi praticamente o mesmo da rede com dez variáveis de entrada e os valores calculados se concentraram em quatro regiões: 1%, 2.5%, 3.8% e 5.4% (Figura 38). A rede não fez uma boa aproximação. O critério de parada atendido foi que o número máximo de épocas foi atingido.

Figura 38 - Saída da rede neuronal e o valor esperado de percentual de frações pesadas no propano, em relação à razão de refluxo de topo e produção de propano (R\_R\_P em %), com topologia de 9 unidades na camada de entrada, 30 neurônios na primeira camada oculta, 15 neurônios na segunda camada, 5 na terceira e um neurônio na camada de saída.



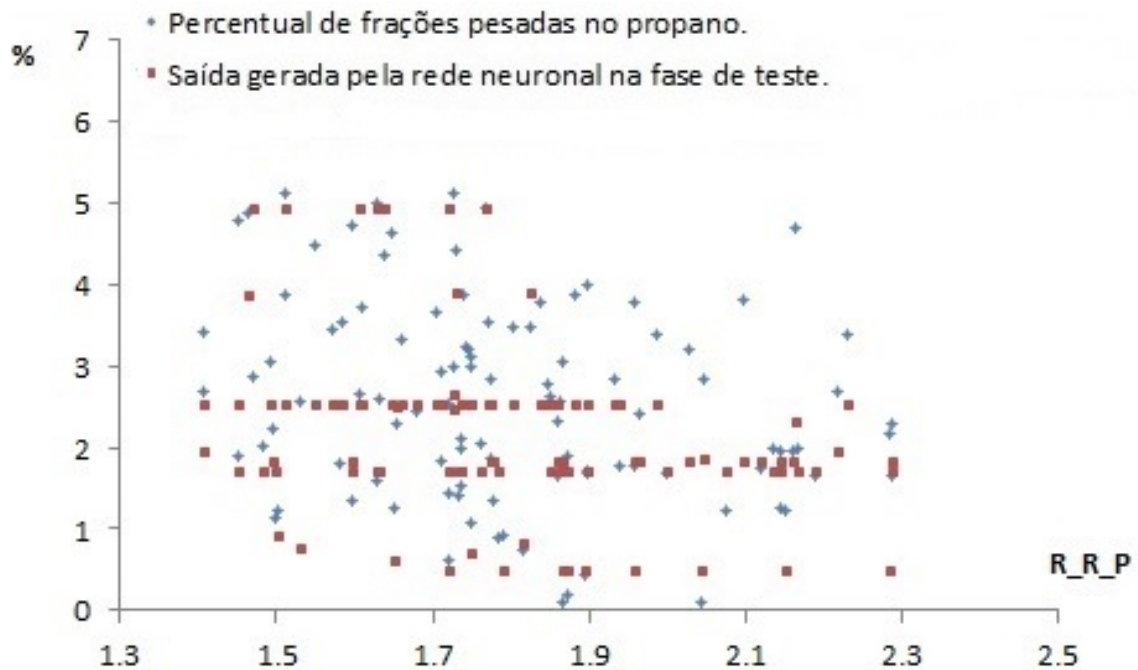
Fonte: O autor, 2017.

#### 4.7.2 Rede com oito unidades de entrada.

Foi retirada a variável de entrada temperatura de fundo (T\_F) da coluna depropanizadora. Este experimento utilizou uma topologia com oito unidades na camada de entrada, trinta neurônios na primeira camada oculta, quinze neurônios na segunda camada, cinco na terceira e um neurônio na camada de saída.

No experimento anterior os menores valores calculados pela rede ficaram próximos de 1. Neste, apesar do erro quadrático médio ter tido um pequeno acréscimo, de 1.01584 para 1.03882 no conjunto da fase de treinamento e de 0.967757 para 1.00604 na de teste, a rede calculou valores próximos de 0.5, fazendo boa aproximação dos menores valores desejados do conjunto de dados, como pode ser observado na Figura 39. Houve uma pequena melhora na aproximação, mas o desempenho da rede ainda foi insatisfatório. O critério de parada atendido foi que o número máximo de épocas foi atingido.

Figura 39 - Saída da rede neuronal e o valor esperado de percentual de frações pesadas no propano, em relação à razão de refluxo de topo e produção de propano (R\_R\_P em %), com topologia de 8 unidades na camada de entrada, 30 neurônios na primeira camada oculta, 15 neurônios na segunda, 5 na terceira e um neurônio na camada de saída.

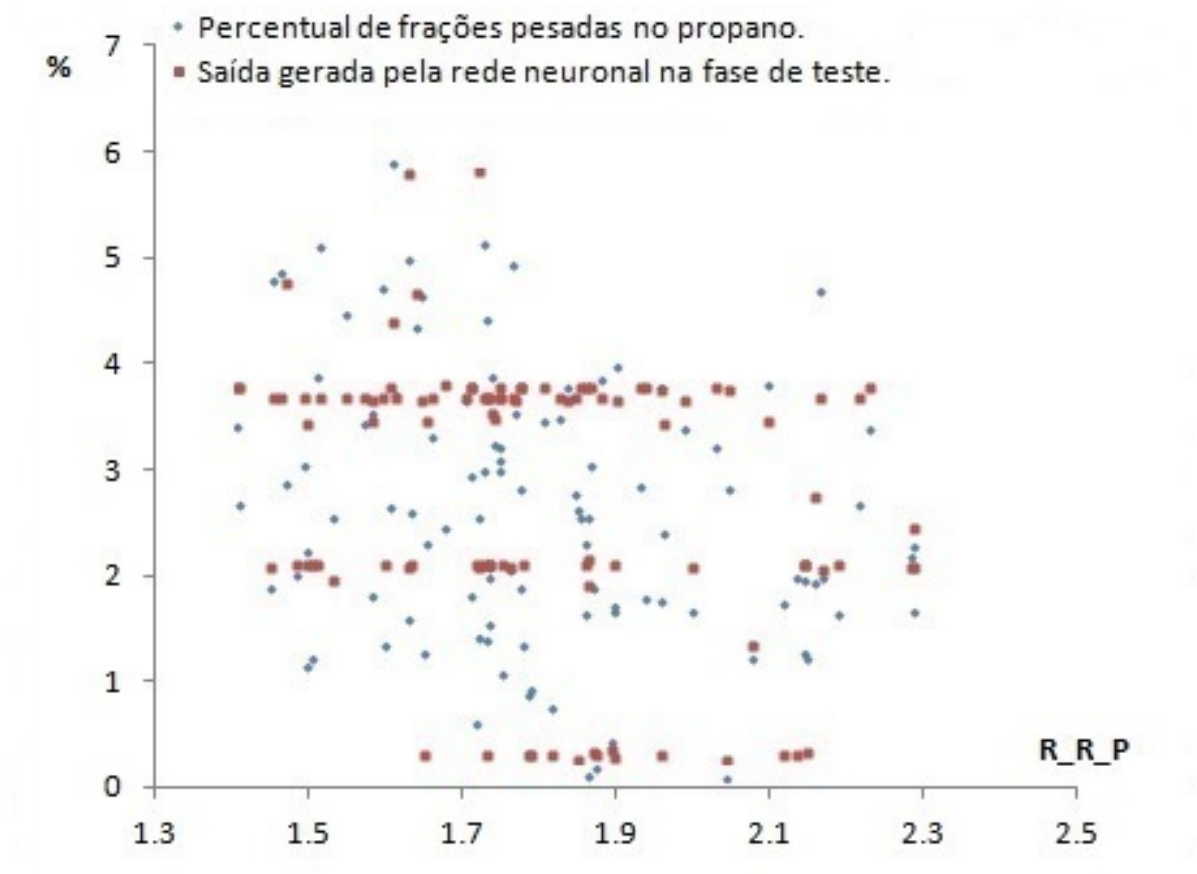


Fonte: O autor, 2017.

#### 4.7.3 Rede com seis unidades de entrada.

Foram retiradas as variáveis de entrada temperatura do prato 25 (T\_P\_25) e temperatura do prato 27 (T\_P\_27) da coluna depropanizadora. Este experimento utilizou uma topologia de seis unidades na camada de entrada, trinta neurônios na primeira camada oculta, quinze neurônios na segunda camada, cinco na terceira e um na camada de saída. O menor erro no conjunto da fase de treinamento foi de 1.04482, e no conjunto da fase de teste de 1.0815. O critério de parada atendido foi que o número máximo de épocas foi atingido e o desempenho ainda ficou abaixo do desejado (Figura 40).

Figura 40 - Saída da rede neuronal e o valor esperado de percentual de frações pesadas no propano, em relação à razão de refluxo de topo e produção de propano (R\_R\_P em %), com topologia de 6 unidades na camada de entrada, 30 neurônios na primeira camada oculta, 15 neurônios na segunda, 5 na terceira e um neurônio na camada de saída.



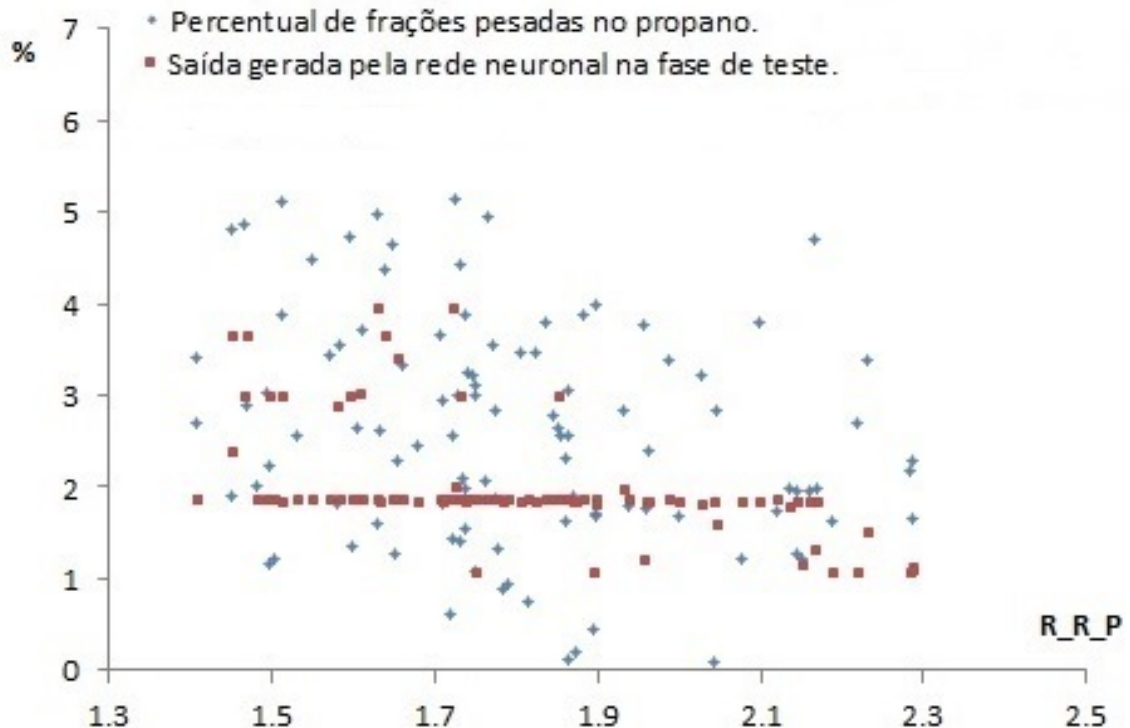
Fonte: O autor, 2017.

#### 4.7.4 Rede com cinco unidades de entrada.

Foi retirada a variável de entrada temperatura de topo (T\_T) da coluna depropanizadora. Este experimento utilizou uma topologia de cinco unidades na camada de entrada, trinta neurônios na primeira camada oculta, quinze neurônios na segunda camada, cinco na terceira e um na camada de saída.

Com a retirada da temperatura de topo (T\_T) da coluna depropanizadora, o erro quadrático médio aumentou, e o menor erro no conjunto da fase de treinamento foi de 1.23665, e no conjunto da fase de teste de 1.32027. A rede neuronal não fez uma boa aproximação (Figura 41). O critério de parada atendido foi que o número máximo de épocas foi atingido.

Figura 41 - Saída da rede neuronal e o valor esperado de percentual de frações pesadas no propano, em relação à razão de refluxo de topo e produção de propano (R.R.P em %), com topologia de 5 unidades na camada de entrada, 30 neurônios na primeira camada oculta, 15 neurônios na segunda camada, 5 na terceira e um neurônio na camada de saída.



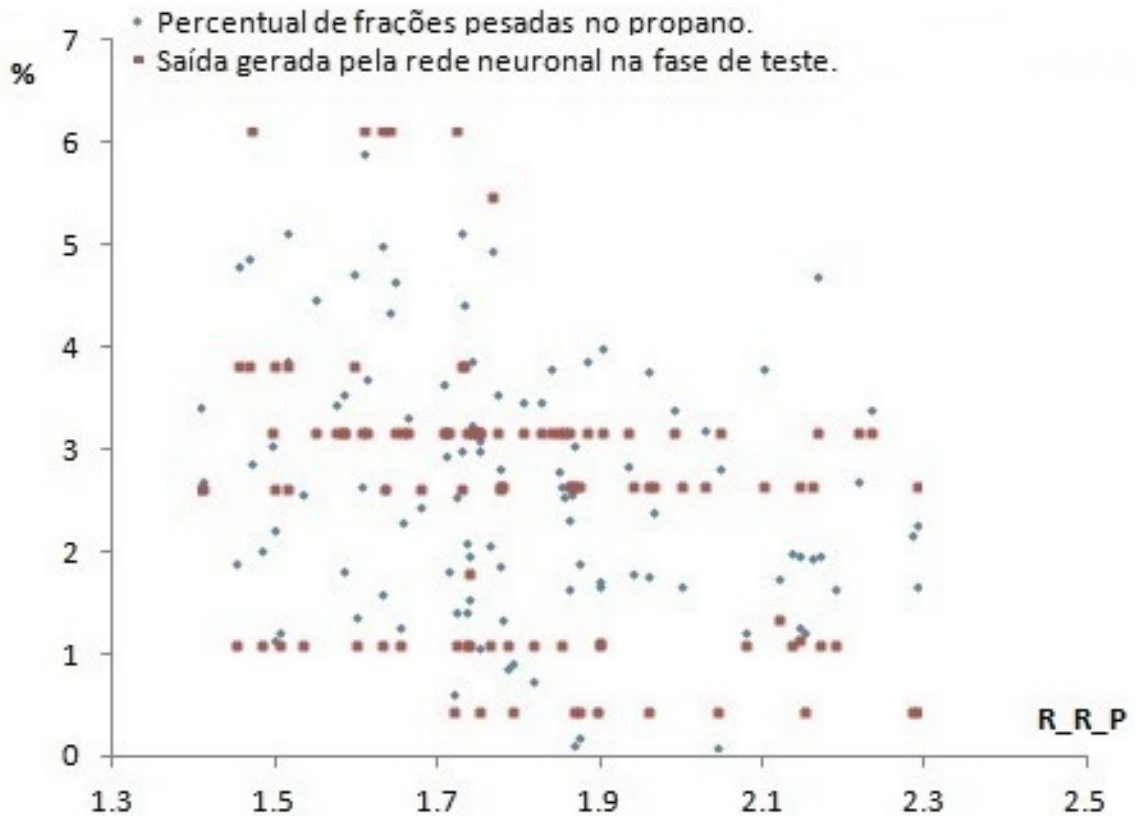
Fonte: O autor, 2017.

#### 4.7.5 Rede após troca de variável de entrada.

A temperatura de topo foi recolocada e foi retirada a variável de entrada temperatura da vazão de topo (T\_V\_T) da coluna depropanizadora. Este experimento utilizou uma topologia de cinco unidades na camada de entrada, trinta neurônios na primeira camada oculta, quinze neurônios na segunda camada, cinco na terceira e um na camada de saída.

O menor erro no conjunto da fase de treinamento foi de 1.01486, e no conjunto da fase de teste de 0.963937. O critério de parada atendido foi que o número máximo de épocas foi atingido.

Figura 42 - Saída da rede neuronal e o valor esperado de percentual de frações pesadas no propano, em relação à razão de refluxo de topo e produção de propano (R\_R\_P em %).



Fonte: O autor, 2017.

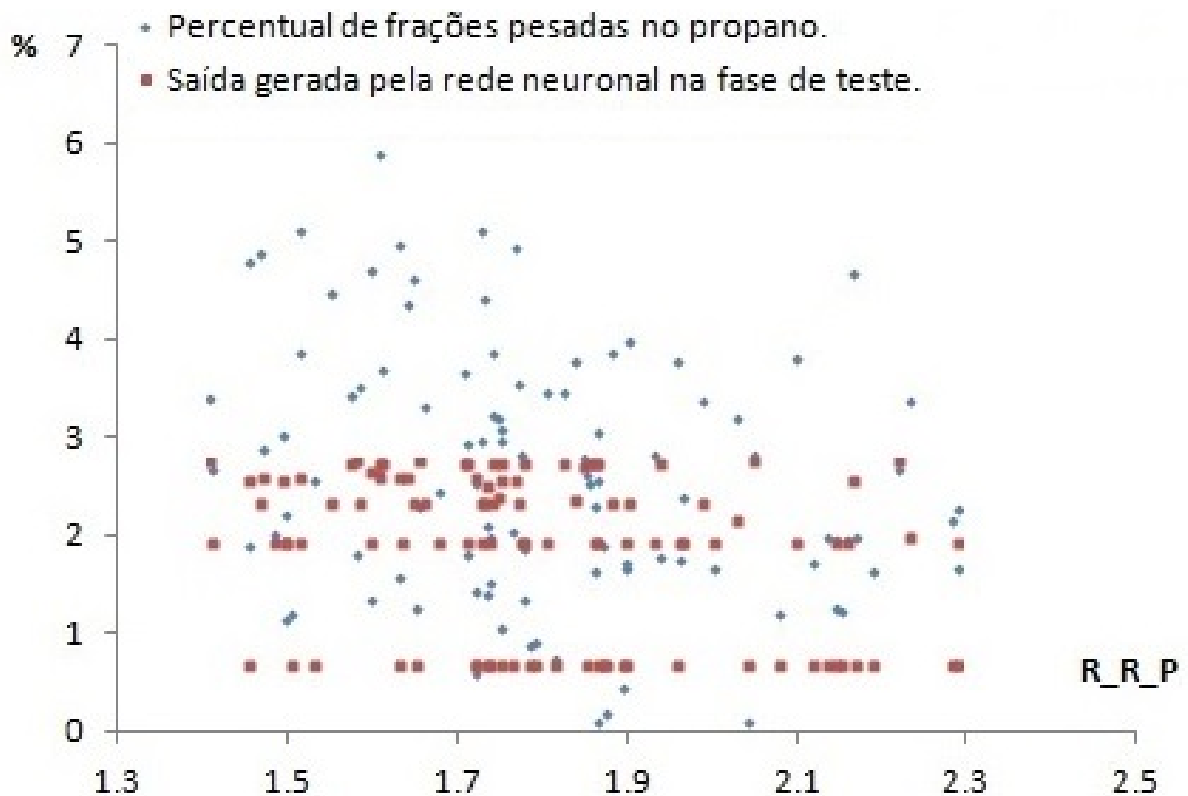
Apesar de o erro quadrático médio ter reduzido e a rede ter melhorado a aproximação, os valores calculados aparecem, em sua maioria, concentrados em torno de 1% e 3% de frações pesadas, conforme pode ser visto na Figura 42.

#### 4.7.6 Rede com seis unidades de entrada, após inclusão de variáveis

As diferenças entre as temperaturas dos pratos, ao longo da coluna, indicam a eficiência no processo de separação, por isso, foi retirada a temperatura do prato vinte e um (T.P.21) e acrescentadas as diferenças de temperatura entre o prato vinte e um (T.P.21) e o prato vinte e cinco (T.P.25) e a diferença entre o prato vinte e um e o prato vinte e sete (T.P.27).

Este experimento utilizou uma topologia de seis unidades na camada de entrada, trinta neurônios na primeira camada oculta, quinze neurônios na segunda camada, cinco neurônios na terceira e um neurônio na camada de saída. O menor erro no conjunto da fase de treinamento foi de 1.06668, e no conjunto da fase de teste de 1.21836. O critério de parada atendido foi que o número máximo de épocas foi atingido.

Figura 43 - Saída da rede neuronal e o valor esperado de percentual de frações pesadas no propano, em relação à razão de refluxo de topo e produção de propano (R.R.P em %), com topologia de 6 unidades na camada de entrada, 30 neurônios na primeira camada oculta, 15 neurônios na segunda camada, 5 na terceira e um neurônio na camada de saída.



Fonte: O autor, 2017.

Esta mudança de variável de entrada fez com que a rede piorasse o desempenho, como mostra a Figura 43.

## 4.8 Experimentos com dados próximos da média

A rede, para gerar bons resultados, necessita de bons dados de treinamento, que representem significativamente o sistema físico. Como em várias combinações de variáveis de entrada e topologia não se conseguiu desempenho aceitável, os dados de entrada receberam novo tratamento.

### 4.8.1 Novo tratamento nos dados de entrada

O novo intervalo das variáveis de entrada foi calculado em função de seu valor médio e do seu desvio padrão, levando-se em consideração que o valor médio deve estar bem próximo do valor normal de operação. O novo intervalo foi calculado de acordo com

$$ValorMáximo = \mu + 1.5 \cdot \sigma \quad (11)$$

e

$$ValorMínimo = \mu - 1.5 \cdot \sigma, \quad (12)$$

onde  $\mu$  é a média aritmética e  $\sigma$  é o desvio padrão.

### 4.8.2 Variáveis de entrada do melhor experimento

As variáveis de entrada foram selecionadas de acordo com o experimento de melhor desempenho e estão apresentadas na Tabela 7.

Tabela 7 - Variáveis de entrada do melhor experimento.

Sigla	Descrição	Intervalo	Unidade
T_T	Temperatura de Topo da Coluna	36.77 – 45.73	°C.
T_P_21	Temperatura do Prato 21	56.64 – 81.67	°C
V_S	Vazão de Vapor	2.380 – 6.510	t/h
P_T	Pressão de Topo	14.61. – 16.98	kgf/cm <sup>2</sup>
R.R.P	Razão Refluxo de Topo e Produção de Propano	0.880 – 3.080	–
%PesadoProp	Percentual de Frações Pesadas no Propano	0.000 – 5.590	%

Fonte: O autor, 2017.

Nestes experimentos o menor erro no treinamento foi de 0.816712, e o menor

erro no teste foi de 0.989814. O critério de parada atendido foi que o número máximo de épocas foi atingido. Houve pequena melhora do desempenho da rede, comprovada pela redução do erro quadrático médio, mas este ainda ficou abaixo do desempenho desejável.



## CONCLUSÕES

A produção de derivados de petróleo requer um controle rigoroso dos parâmetros de qualidade. Identificar, no menor tempo possível, que a produção está fora dos limites das especificações de qualidade representa um grande ganho econômico.

Sensores virtuais são muito utilizados para identificar se a produção está fora dos limites de especificação, suprindo em parte, o problema gerado pelo longo tempo necessário para obter os resultados dos ensaios feitos em laboratório. O que fundamenta a utilização destes sensores virtuais é o fato dos parâmetros de qualidade dos derivados serem dependentes de variáveis do processo, tais como vazão, temperatura e pressão.

Uma vez que a relação entre os parâmetros de qualidade e as variáveis de processo geralmente é não linear, a abordagem de rede neuronal é uma opção conveniente para a sua modelagem. Além disso, é razoável supor que há uma dependência contínua dos parâmetros de qualidade com relação às variáveis de processo. Deste modo podemos explorar, também, a capacidade da rede neuronal de aproximar qualquer função contínua.

Uma grande dificuldade na modelagem através de rede neuronal, com arquitetura de perceptron de múltiplas camadas, é a escolha da topologia. As heurísticas, como Kolmogorov e Fletcher-Gloss, normalmente são aplicáveis em redes com somente uma camada oculta e para alguns tipos de problemas específicos, além de não considerarem a quantidade de dados, sua disposição no espaço amostral e a complexidade do problema.

A quantidade de unidades da camada de entrada e dos neurônios da camada de saída é determinada em função das características do problema. Nos experimentos preliminares, com as funções trigonométricas, tivemos uma unidade na camada de entrada e um neurônio na camada de saída. Para este tipo de problema, aproximar as funções seno e cosseno, os valores desejados da saída estão entre o intervalo  $[-1, 1]$ , devido à imagem destas funções. Então, com uma topologia de uma camada oculta, o aumento dos neurônios desta camada, pode levar a saturação da função de ativação do neurônio da saída, em virtude do conseqüente aumento do número de entradas deste neurônio. Quando este limite de neurônios é atingido, a solução é acrescentar mais uma camada oculta.

Observou-se, também, nos testes preliminares, que o algoritmo de retropropagação se mostrou bastante sensível ao intervalo dos valores da entrada. Quando estes valores, em termos absolutos, foram superiores a 8.5 tivemos problemas de saturação da função de ativação, o que refletiu em saída calculada pela rede neuronal, praticamente, constante. As variáveis de entrada da rede neuronal, para o problema da coluna depropanizadora, possuem valores superiores, em relação ao do problema da função trigonométrica, na ordem de  $10^2$  e  $10^3$ .

As redes neuronais têm sua eficiência fortemente relacionada à precisão das

variáveis de entrada, quer seja nas fases de treinamento e teste como na de execução, por isso, os medidores destas variáveis devem estar sempre bem ajustados. Em processos complexos, como os das refinarias, identificar problemas de medição, tão logo eles apareçam, nem sempre é possível, desta forma, esta modelagem pode não ser tão simples.

Além dos problemas de medição, que podem estar presentes nos dados utilizados durante nossos experimentos, existem as variáveis que têm influência em colunas de destilação e não são medidas. No caso da depropanizadora temos a alteração na composição da carga, que caso não seja perceptível através das variáveis medidas, somente é descoberta nos ensaios de laboratório.

Mesmo que a vazão de vapor (V\_S) não varie, sua temperatura, que não é medida, pode variar, acarretando alterações na temperatura de retorno do produto de fundo, após este ser aquecido. Somente quando o perfil de temperatura da torre (temperatura dos pratos) alterar é possível identificar alteração na coluna, em função do aumento da temperatura do vapor. Neste momento, os produtos, propano e butano, já poderiam estar fora dos limites de especificação. Estas variáveis, que não possuem medição, podem ter comprometido o treinamento da rede, caso elas tenham variado durante a coleta dos dados, que foram utilizados em nossos experimentos.

Outro ponto, não levado em consideração nos nossos experimentos, foi a variável tempo. É de se esperar que variações na temperatura de topo acarretem alterações nas frações pesadas na produção de propano (topo) muito mais rapidamente do que variações na vazão de vapor, que aquece o fundo da coluna depropanizadora.

A incorporação da variável tempo, numa rede neuronal, pode ser feita na amostragem das variáveis de entrada da seguinte forma

$$x_0 = (t_0, t_1, t_2, \dots, x_n) \quad (13)$$

$$x_1 = (t_2, t_3, t_4, \dots, t_{n+2}) \quad (14)$$

$$x_2 = (t_1, t_2, t_3, \dots, t_{n+1}), \quad (15)$$

representando três séries temporais com intervalos constantes entre os tempos discretos. As diferenças de tempo entre  $t_0$  e  $t_1$  e  $t_0$  e  $t_2$  podem ser levantadas utilizando técnicas de identificação de sistemas, que é a área de modelagem matemática que estuda técnicas alternativas a modelagem pela física do processo ou modelagem conceitual, utilizando algoritmos para obter (identificar) modelos de sistemas dinâmicos a partir de dados gerados pelo próprio sistema (AGUIRRE, 2007).

Uma rede neuronal de múltiplas camadas, com o algoritmo de retropropagação, pode ser considerada uma ferramenta para realizar um mapeamento não linear de entrada-saída de natureza geral (HAYKIN, 2003). Para determinar se esta rede está fazendo este mapeamento com precisão, levanta-se o seguinte questionamento:

Qual é o número mínimo de camadas ocultas em um perceptron de múltiplas

camadas com um mapeamento de entrada-saída que fornece uma realização aproximada de qualquer mapeamento contínuo? (HAYKIN, 2003)

O teorema da aproximação universal responde esta pergunta com somente uma camada oculta. Entretanto, o teorema não diz que a única camada oculta é ótima no sentido do tempo de aprendizagem, facilidade de implementação, ou generalização (HAYKIN, 2003), propriedade fundamental das redes neuronais. Em relação à generalização, nos experimentos da coluna depropanizadora, a rede apresentou deficiência, caracterizada pelo erro quadrático médio maior nos testes do que nos treinamentos.

Devido ao número parâmetros configuráveis de uma rede do tipo perceptron de múltiplas camadas (número de camadas ocultas, número de neurônios de cada camada oculta, tipo de função de ativação), esta modelagem pode ser bastante complexa. Outra dificuldade pode ser a escolha das variáveis de entrada, que além de influenciar na escolha da função de ativação, podem comprometer o treinamento, por exemplo, se tivermos entradas redundantes, isto é, que contenham o mesmo tipo de informação em relação às saídas da rede.

## Recomendações

Nosso conjunto de dados, na forma de matriz  $M_{LC}$ , sendo  $L$  o número de linhas e  $C$  o número de colunas, com a retirada dos dados espúrios, seção 4.1, teve sua dimensão  $L$  reduzida. Após a decisão de calcular em redes distintas as duas especificações, fixar o intervalo da variável de entrada  $R_{R-P}$  entre [1.4, 2.3] e retirar os registros em que o percentual de frações pesadas no propano (%PesadoProp) fossem maiores que 6, a matriz do conjunto de dados foi reduzida nas dimensões  $L$  e  $C$ , subseção 4.3.3. Com a retirada das variáveis de entrada, seção 4.7, para validar a relevância de cada variável no cálculo da especificação desejada, a dimensão  $C$  foi reduzida gradualmente até chegar a 5 no experimento de melhor desempenho. Esta ordem de tratar os dados e reduzir o intervalo das variáveis de entrada e, somente, posteriormente reduzir a dimensionalidade do problema, através da retirada de variáveis de entrada, não nos pareceu a mais adequada. Esta é, em nossa opinião, a melhor ordem:

- Retirada dos dados espúrios.
- Verificação da relevância de cada variável de entrada, com a possível redução da dimensão  $C$ .
- Baseado nas características do problema que se tenta resolver, reduzir o intervalo das variáveis de entrada, quando possível, reduzindo a dimensão  $L$ .

O perfeito conhecimento dos dados para treinamento é fundamental na modelagem de rede neuronal. Analisar as variáveis de entrada, com o objetivo de reduzir a dimensão do problema, possibilitando eliminar redundâncias e a escolha adequada da função de ativação para o tipo de problema, pode facilitar o tratamento dos dados e o treinamento da rede neuronal.

A normalização utilizada em nossos experimentos não evitou, em alguns casos, a saturação das funções de ativação, portanto pode-se utilizar a inicialização dos pesos sinápticos sugerida por LeCun (LECUN, 1989), onde os valores devem estar compreendidos entre

$$\frac{-2.4}{N} \leq W_{ij}^0 \leq \frac{2.4}{N}, \quad (16)$$

sendo  $N$  o número de unidades de entrada da rede neuronal e  $W_{ij}^0$  a matriz de pesos sinápticos que conectam as unidades da camada de entrada aos neurônios da primeira camada oculta.

## Trabalhos futuros

Durante o desenvolvimento deste trabalho foi desenvolvido um sensor virtual (inferência) para a torre depropanizadora, que tem apresentado bom desempenho. De qualquer forma, existem especificações de produtos que, ainda, demandam a criação destas inferências, sensores virtuais, pelos mesmos motivos expostos neste trabalho, em relação às especificações do propano e do butano.

Uma alternativa, para a modelagem, são as redes de funções de base radial (RBF), que possuem somente uma camada oculta, portanto mais simples de serem configuradas e para problemas de aproximação de funções, as redes RBF são especialmente recomendadas, em função de superfícies com picos e vales regulares, uma vez que possibilitam a obtenção de modelos precisos e eficientes (XIE; YU; BOGDAN, 2011).

De acordo com (GOMES, 2010) funções de ativações diferentes das tradicionais, logística e tangente hiperbólica, podem ser utilizadas, já que estas possuem a capacidade de melhorar o desempenho da rede perceptron de múltiplas camadas e diminuir sua complexidade.

## REFERÊNCIAS

- AGUIRRE, L. A. *Introdução à identificação de sistemas: Técnicas lineares e não lineares aplicadas a sistemas reais*. Belo Horizonte: UFMG, 2007.
- CAMARGO, P. R. C. *Processos de Refino*. Rio de Janeiro, 2001.
- CYBENKO, G. Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals, and Systems*, v. 2, p. 303–314, 1989.
- FACURE, M. *Funções de Ativação*. 2017. Disponível em: <https://matheusfacure.github.io/2017/07/12/activ-func/>. Acesso em: 6 agosto 2017.
- FINOCCHIO, M. A. F. *Noções de Redes Neurais Artificiais*. 2014. 28 p.
- FUNAHASHI, K. *On the approximate realization of continuous mappings by neural networks*. [S.l.]: Neural Networks, 1989. 183-192 p.
- GOMES, G. S. S. *Novas Funções de Ativação: em redes neurais artificiais multilayer perceptron*. 2010. 159 f. Tese (Doutorado em Ciência da Computação) — Centro de Informática, Universidade Federal de Pernambuco, Recife, 2010.
- HAYKIN, S. *Neural Networks: A Comprehensive Foundation*.: para engenharia e ciências aplicadas. 2. ed. Nova Deli: Pearson Education, 2003. 183-200 p.
- HERTZ, J. A.; PALMER, R. G.; KROGH, A. S. *Introduction to the theory of neural computation*. California: Addison-Wesley Publishing Company, 1991. 1-130 p.
- HODGKIN, A. L.; HUXLEY, A. F. *A quantitative description of membrane current and its application to conduction and excitation in nerve*. Londres: WILEY, 1952.
- LECUN, Y. *Generalization and network design strategies*. Toronto, 1989.
- PADUA, A. et al. *Redes Neurais Artificiais - Teoria e Aplicações*. Rio de Janeiro: LTC, 2000.
- PLATE, T. *Neural Network: tricks of the trade*. New York: Springer, 1988.
- SILVA, I. N.; SPATTI, D. H.; FLAUZINO, R. A. *Redes neurais artificiais:: para engenharia e ciências aplicadas*. 1. ed. São Paulo: Artiber, 2010. 29 p.
- USP. *Escoamentos do vapor e do líquido em colunas de pratos*. 2016. Disponível em: [http://sites.poli.usp.br/p/luiz.terron/destilacao/2\\_paginas\\_equipamentos/2\\_2\\_distilacao\\_colunas\\_internas.htm](http://sites.poli.usp.br/p/luiz.terron/destilacao/2_paginas_equipamentos/2_2_distilacao_colunas_internas.htm). Acesso em: 16 outubro 2016.
- XIE, T.; YU, H.; BOGDAN, W. Comparison between traditional neural networks and radial basis function networks. *Industrial Electronics (ISIE), 2011 IEEE International Symposium on*, Gdansk, Poland, p. 6, 15 aug. 2011.

## ANEXO A – Pseudocódigo

Este é o pseudocódigo de uma rede perceptron de múltiplas camadas com duas entradas, duas camadas ocultas e um neurônio na camada de saída. Neste caso tem-se subjacente à rede uma função real de duas variáveis.

### A.0.1 Treinamento.

Cada treinamento realiza um ciclo completo do algoritmo e o conjunto de dados, de cada um, é o mesmo. O treinamento é composto de épocas, limitadas em um número máximo, para quando o algoritmo não converge para um erro, de acordo com a Equação (5), menor que o valor mínimo de erro determinado.

Notação:

- O escalar  $t$  refere-se ao número de treinamentos.
- O escalar  $p$  refere-se ao número de padrões apresentados a rede.
- $\eta$  é taxa de aprendizagem e varia entre 0.1 e 1.
- O escalar  $n$  refere-se ao número de unidades de entrada da rede. O escalar  $n_1$  refere-se ao número de neurônios da primeira camada oculta. O escalar  $n_2$  refere-se ao número de neurônios da segunda camada oculta. O escalar  $n_3$  refere-se ao número de neurônios da camada de saída.
- $W_{ji}^{(c)}$  são matrizes de pesos sinápticos conectando o neurônio  $j$  da camada  $(c)$  ao neurônio  $i$  da camada  $(c - 1)$
- $d_j^{(l)}$  é o valor desejado para o neurônio  $j$  da camada de saída, na iteração  $l$ .
- $Y_j^{(c)}$  é a saída do neurônio  $j$  da camada  $c$ .
- O vetor  $v$  contém índices em ordem aleatória para apresentar os padrões de treinamento em diferentes ordenações em cada época.

Algoritmo 1 - Backpropagation.

1. Criar rede conforme topologia escolhida
  2. Obter conjunto de amostras de treinamento  $x^{(p)}$
  3. Obter conjunto saídas desejadas  $d^{(p)}$
  4. Associar cada valor saída desejada  $d_i$  respectivas entradas  $x_i$ .
  5. Definir taxa aprendizagem e precisão requerida.
  6. **para**  $k$  de 1 até  $t$ , **fazer** ..... {para cada possível valor de 'k'.}
  7. | Inicializar matriz de pesos  $W$  com valores aleatórios intervalo  $[-1, 1]$
  8. | Iniciar contador número épocas
  9. | Inicializar vetor  $v$  com valores de  $[1, p]$  com ordem aleatória
  10. | para apresentar conjunto de dados em diferentes ordens em cada época.
  11. | **enquanto** ( $E_{medio} > 0.01$  e épocas  $< 10000$  e Contador  $< 5$ ), **fazer**
  12. | | **para**  $l$  de 1 até  $p$ , **fazer** ..... {para todas as amostras de treinamento}
  - {j variando de 1 até  $n_1$ }
  13. | | | Obter  $I_j^{(1)}$  e  $Y_j^{(1)}$  conforme :
  14. | | |  $I_j^{(1)} = \sum_{i=0}^n W_{ji}^{(1)} \cdot x_i(v(l))$
  15. | | |  $Y_j^{(1)} = g(I_j^{(1)})$
  - {j variando de 1 até  $n_2$ }
  16. | | | Obter  $I_j^{(2)}$  e  $Y_j^{(2)}$  conforme:
  17. | | |  $I_j^{(2)} = \sum_{i=0}^{n_1} W_{ji}^{(2)} \cdot Y_i^{(1)}$
  18. | | |  $Y_j^{(2)} = g(I_j^{(2)})$
  - {j variando de 1 até  $n_3$ }
  19. | | | Obter  $I_j^{(3)}$  e  $Y_j^{(3)}$  conforme:
  20. | | |  $I_j^{(3)} = \sum_{i=0}^{n_2} W_{ji}^{(3)} \cdot Y_i^{(2)}$
  21. | | |  $Y_j^{(3)} = g(I_j^{(3)})$
  22. | | | Determinar  $\delta_j^{(3)}$  conforme
  23. | | |  $e_j(l) = d_j(v(l)) - Y_j^{(3)}(l)$
  24. | | |  $\delta_j^{(3)} = (e_j(l)) \cdot g'(I_j^{(3)})$
  - {Ajustar pesos sinápticos dos neurônios de camada de saída  $c = 3$ }
  25. | | | Ajustar  $W_{ji}^{(3)}$  conforme
  26. | | |  $W^{(3)ji} \leftarrow W_{ji}^{(3)} + \eta \cdot \delta_j^{(3)} \cdot Y_i^{(2)}$
  27. | | | Determinar  $\delta_j^{(2)}$  conforme
  28. | | |  $\delta_j^{(2)} = (\sum_{k=0}^{n_3} \delta_k^{(3)} \cdot W_{kj}^{(3)}) \cdot g'(I_j^{(2)})$
  - {Ajustar pesos sinápticos neurônios segunda camada oculta  $c = 2$ }
  29. | | | Ajustar  $W_{ji}^{(2)}$  conforme
  30. | | |  $W^{(2)ji} \leftarrow W_{ji}^{(2)} + \eta \cdot \delta_j^{(2)} \cdot Y_i^{(1)}$
  31. | | | Determinar  $\delta_j^{(1)}$  conforme
  32. | | |  $\delta_j^{(1)} = (\sum_{k=0}^{n_2} \delta_k^{(2)} \cdot W_{kj}^{(2)}) \cdot g'(I_j^{(1)})$
- continua —

Algoritmo 1 - Backpropagation. (Continuação)

— continuação —

{Ajustar pesos sinápticos neurônios primeira camada oculta  $c = 1$ }

```

33. | | | Ajustar  $W_{ji}^{(1)}$  conforme
34. | | |  $W^{(1)ji} \leftarrow W^{(1)ji} + \eta \cdot \delta_j^{(1)} \cdot x_i(v(l))$ 
35. | | | Obter  $Y_i^{(3)}$  conforme (14, 15, 17, 18, 20 e 21)
36. | | | Calcular erro( $Em$ )
37. | | |  $EmAtual \leftarrow Em$ 
38. | | | incrementar época
39. | | fim para
40. | | se ( $\text{abs}(E_{medioanterior} - E_{medio}) < 0.1$ ), então . {erro não está reduzindo ...}
41. | | | incrementar Contador
42. | | senão
43. | | |  $Contador \leftarrow 0$ 
44. | | fim se
45. | fim enquanto
46. fim para
    incrementar Contador
senão
     $Contador \leftarrow 0$ 
fim se
fim enquanto
fim para

```