

Relatório 06

Redes Neurais Artificiais

Multi Layer Perceptron

Cristiano Lopes Moreira

Matrícula: 119103-0

Aluno		RA/Matrícula		Professor		Tipo	
Cristiano Lopes Moreira		119103-0		Dr Reinaldo Bianchi		Relatório de implementação	
Data	Versão	Turma		Nome do arquivo			Página
12/12/2019	1	2º. Semestre de 2019		PEL_208_Relatório_06_Cristiano_Moreira.doc			1 (32)

Sumário

1.	Introdução	3
2.	Desenvolvimento teórico	4
3.	Proposta de implementação Algoritmo Multi Layer Perceptron	8
4.	Experimentação e Resultados	9
4.1.	Dados Porta Logica XOR	10
4.2.	Iris Fisher	11
4.3.	Seed	16
4.4.	Breast Cancer Wisconsin (Original) Data Set	23
5.	Conclusão	32
6.	Referências	32

Aluno		RA/Matrícula	Professor	Tipo	
Cristiano Lopes Moreira		119103-0	Dr Reinaldo Bianchi	Relatório de implementação	
Data	Versão	Turma	Nome do arquivo		Página
12/12/2019	1	2º. Semestre de 2019	PEL_208_Relatório_06_Cristiano_Moreira.doc		2 (32)

1. Introdução

"Cybernetics and Romanticism

Perceptrons have been widely publicized as "pattern recognition" our "learning" machines and as such have been discussed in a large number of books, journal articles and voluminous "reports." Much of this writing (some exceptions are mentioned in our bibliography) is without scientific value and we will not usually refer by name to the works we criticize. The sciences of computation and cybernetics began, and it seems quite rightly so, with a certain flourish of romanticisms. They were laden with attractive and exciting new ideas which have already borne rich fruit. Heavy demands of rigor and cautions could have held this development to a much slower pace; only the future could tell which directions were to be the best. We feel, in fact, that the solemn experts who most complained about the "exaggerated claims" of the cybernetic enthusiasts were, in the balance much more in the wrong. But now the time has come for maturity, and this requires us to match our speculative enterprise with equally imaginative standards of criticism." (MINSKY; PAPERT, 1988)

A busca, pela observação a neurociência, de um modelamento matemático das atividades sinápticas do cérebro humano, em especial das células chamadas neurônios, auxiliou na construção de técnicas voltadas à ideia de uma inteligência artificial (IA) inspirada em hipóteses de criar uma rede neural artificial (RNA) pela modelagem de um neurônio, o perceptron (ROSENBLATT 1958) como um classificador binário linear. Essa ideia foi classificada por Minsky e Papert (1988) como um romantismo cibernético, mesmo admitindo a evolução em várias áreas da ciência pelo uso de sistemas lineares "In fact, we feel that the critical advances in many branches of science and mathematics began with good formulations of the "linear" systems, and these machines are our candidate for beginning the study of "parallel machines" in general." (MINSKY; PAPERT, 1988);

O livro publicado por Minsky e Papert trouxe afirmações sobre a impossibilidade do modelo do perceptron de realizar aprendizado distinto de uma simples regressão linear por não ser possível, de acordo com os autores, transferir conhecimento para as camadas ocultas em uma rede de múltiplas camadas de perceptrons. Como afirmado por Minsky e Papert, o erro entre a saída esperada e a saída estimada é objetivo e observável na camada de saída da rede, porém, os erros nas camadas ocultas são desconhecidos uma vez que os dados de treinamento não

Aluno		RA/Matrícula	Professor	Tipo	
Cristiano Lopes Moreira		119103-0	Dr Reinaldo Bianchi	Relatório de implementação	
Data	Versão	Turma	Nome do arquivo		Página
12/12/2019	1	2º. Semestre de 2019	PEL_208_Relatório_06_Cristiano_Moreira.doc		3 (32)

informam quais valores os elementos ocultos devem ter, o que gera uma incerteza sobre o modelo e um desafio científico sobre o tema.

O desenvolvimento do algoritmo Back-Propagation (RUMELHART, HILTON, E WILLIAMS, 1986), pela derivação do gradiente do erro total, propõem que para ser possível as camadas ocultas de uma rede neural artificial armazenarem conhecimento basta conhecer e propagar o erro da saída à entrada da RNA pelo gradiente do erro proporcional a cada peso e a cada elemento oculto do rede. “Os erros dos elementos processadores da camada de saída (conhecidos pelo treinamento supervisionado) são retro propagados para as camadas intermediárias.”

A descoberta de Rumelart, Hilton e Williams proporcionou o retorno das pesquisas em aprendizado de máquina que culminaram em diversas aplicações nas áreas: reconhecimento e classificação de padrões de escrita e fala; reconhecimento de faces em visão computacional; controle e previsão de ações no mercado financeiro; identificação de anomalias em imagens médicas; entre outras.

O objetivo deste trabalho é implementar uma rede neural multi layer perceptron com back-propagation para classificação de elementos em uma base de dados conhecida e verificar a eficiência de reconhecimento e agrupamento do algoritmo.

2. Desenvolvimento teórico

A regressão linear é um dos algoritmos mais conhecidos e utilizados em estatística e em aprendizado de máquina, sua representação linear é uma equação matemática que melhor descreve uma reta que se encaixa entre os pontos amostrados de uma variável X com saída representando a descrição de um fenômeno.

$$y = f(X) = \alpha + \beta X \quad (1)$$

Aluno Cristiano Lopes Moreira		RA/Matrícula 119103-0	Professor Dr Reinaldo Bianchi	Tipo Relatório de implementação	
Data 12/12/2019	Versão 1	Turma 2º. Semestre de 2019	Nome do arquivo PEL_208_Relatório_06_Cristiano_Moreira.doc		Página 4 (32)

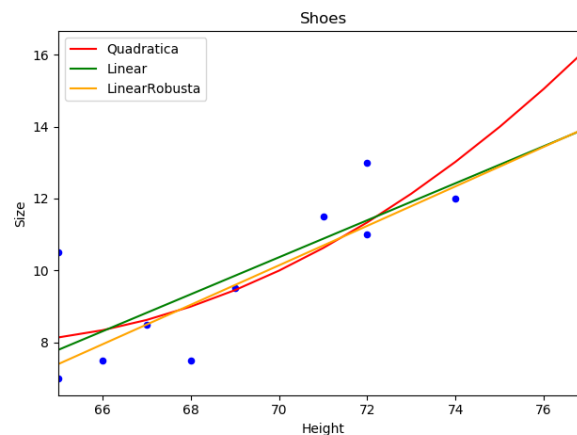


Fig1: regressão linear.

Suponha que Y seja conhecido até um número finito de pontos p dos parâmetros $\beta = (\beta_1, \dots, \beta_p)$, ou seja, $Y = X\beta$. Estimamos β pelo valor de coeficiente $\hat{\beta}$ que melhor se ajusta aos dados.

A rede neural artificial perceptron, assim como o método de regressão linear, busca encontrar os coeficientes da reta que melhor representa estes pontos, porém, distinto do método de regressão linear pelos mínimos quadrados, o coeficiente de estimativa calculado pela RNA perceptron é obtido através de diversas interações, chamadas de treinamento ou aprendizado, que minimizam o erro do coeficiente com base na confirmação do resultado esperado balizado por um coeficiente de aprendizado. Uma vez concluído o processo de treinamento, o perceptron pode ser utilizado para estimar casos, que não fazem parte do conjunto de dados do treinamento, através da somatória de todas as variáveis de entrada, balizadas pelos pesos sinápticos, e pela verificação pela função de ativação se as entradas alcançam seu nível de saturação para acionar a saída do perceptron com 1 ou -1 (classificação binária), que ao final é a estimativa de qual classe pertence o item pesquisado

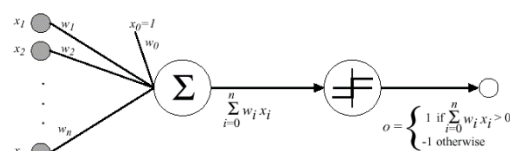


Fig2: perceptron (HASTIE, 2008).

Aluno		RA/Matrícula		Professor	Tipo
Cristiano Lopes Moreira		119103-0		Dr Reinaldo Bianchi	Relatório de implementação
Data	Versão	Turma		Nome do arquivo	Página
12/12/2019	1	2º. Semestre de 2019		PEL_208_Relatório_06_Cristiano_Moreira.doc	5 (32)

Matematicamente os coeficientes da reta, ou pesos sinápticos, são vetorialmente representados por “w”, com seu aprendizado/atualização, realizado pelas diversas interações nos ciclos de aprendizado (episódios ou épocas) com as equações abaixo:

$$w_i = w_i + \Delta w_i \quad (2)$$

Sendo: $\Delta w_i = \eta(t - o)x_i \quad (3)$

Onde: t é o valor desejado (4)
 o é a saída do perceptron
 x_i é a entrada do perceptron
 η é a taxa de aprendizado

Função de ativação
$$f(W^T X) = \begin{cases} 1 & \text{se } \sum_{i=0}^n w_i x_i > 0 \\ -1 & \text{caso contrário} \end{cases} \quad (5)$$

Parâmetro	Variável Representativa	Tipo Característico
Entradas	X_i (i-ésima entrada)	Reais ou binários
Pesos Sinápticos	w_i (associado a x_i)	Reais iniciados aleatoriamente
Saída	y	Binária
Função de Ativação	$g(x)$	Degrau bipolar
Processo de Treinamento	-----	Supervisionado
Regra de Aprendizado	-----	Regra de Hebb

Multi Layer Perceptron é uma rede de múltiplas camadas com conexões progressivas (feedforward) sem realimentação, um arranjo de um conjunto de perceptrons em diversas camadas, com no mínimo uma camada oculta, que possibilita pelo agrupamento/soma de diversas redes lineares em um modelo computacional de classificação não linear.

Aluno		RA/Matrícula		Professor		Tipo	
Cristiano Lopes Moreira		119103-0		Dr Reinaldo Bianchi		Relatório de implementação	
Data	Versão	Turma		Nome do arquivo			Página
12/12/2019	1	2º. Semestre de 2019		PEL_208_Relatório_06_Cristiano_Moreira.doc			6 (32)

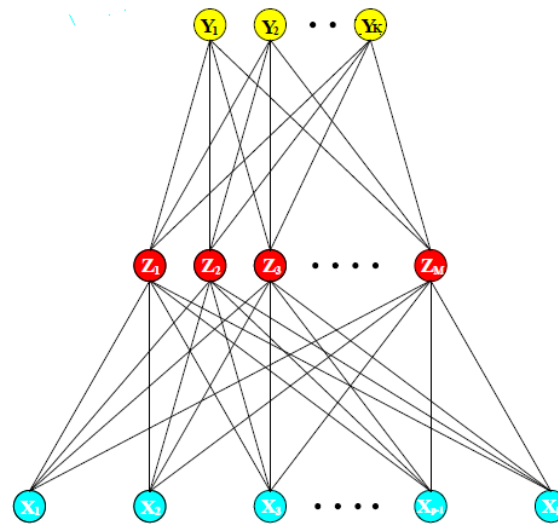


Fig3: Schematic of a single hidden layer, feed-forward neural network (HASTIE, 2008).

Sua atualização/aprendizado é realizada pelo algoritmo back-propagation, que propaga o erro da saída aos elementos escondidos.

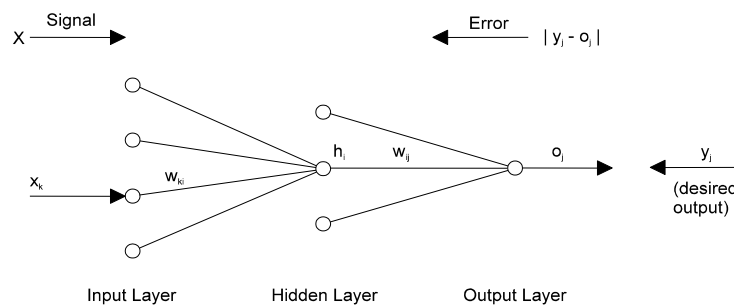


Fig4: back-propagation (HASTIE, 2008).

Sendo:

$$\Delta w_{ij} = \eta \left(\sum_k e_k w_{jk} \right) f'(net_j) s_i \quad (6)$$

Onde:

t é o valor desejado (4)

o é a saída do perceptron

x_i é a entrada do perceptron

η é a taxa de aprendizado

Aluno		RA/Matrícula		Professor		Tipo	
Cristiano Lopes Moreira		119103-0		Dr Reinaldo Bianchi		Relatório de implementação	
Data	Versão	Turma		Nome do arquivo			Página
12/12/2019	1	2º. Semestre de 2019		PEL_208_Relatório_06_Cristiano_Moreira.doc			7 (32)

3. Proposta de implementação Algoritmo Multi Layer Perceptron

Pseudocódigos:

Treinamento

Obtém-se um conjunto de dados;

Associar a saída desejada para cada amostra obtida;

Iniciar um vetor w com valores aleatórios pequenos;

Especificar uma taxa de aprendizagem η_{aprend}

Aplica-se um padrão de entrada x com respectivo vetor de saída t

Calcula-se as saídas dos processadores até a camada de saída

Calcula-se o erro para cada processador da camada de saída

Se erro \leq tolerância para todos os processadores então volta ao passo aplica-se um padrão de entrada x com respectivo vetor de saída t

Atualiza os pesos de cada processador começando pela camada de saída até a primeira camada escondida

Aluno		RA/Matrícula	Professor	Tipo	
Cristiano Lopes Moreira		119103-0	Dr Reinaldo Bianchi	Relatório de implementação	
Data	Versão	Turma	Nome do arquivo		Página
12/12/2019	1	2º. Semestre de 2019	PEL_208_Relatório_06_Cristiano_Moreira.doc		8 (32)

4. Experimentação e Resultados

Para verificar o funcionamento do algoritmo do Multi Layer Perceptron, foi realizada a implementação em Python confrontando os resultados entre a classificação indicada na base de dados e o agrupamento proposto pelo algoritmo:

Ambiente:

PyCharm 2019.2.2 (Professional Edition) Build#PY-192.6603.34

Python 3.7.5 (tags/v3.7.5:5c02a39a0b, Oct 15 2019, 01:31:54) on win32

Bibliotecas:

xlrd-1.2.0 (leitura de arquivos do Excel - base de dados)

numpy-1.17.4 (gestão de matrizes)

Base de Dados:

dbTraining.xlsx	XOR	Porta logica XOR de 2 entradas
	Íris	http://archive.ics.uci.edu/ml/datasets/Iris
	Seed	https://archive.ics.uci.edu/ml/datasets/seeds
	Breast Cancer Wisconsin	https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+%28Original%29

Aluno		RA/Matrícula		Professor		Tipo	
Cristiano Lopes Moreira		119103-0		Dr Reinaldo Bianchi		Relatório de implementação	
Data	Versão	Turma		Nome do arquivo			Página
12/12/2019	1	2º. Semestre de 2019		PEL_208_Relatório_06_Cristiano_Moreira.doc			9 (32)

4.1. Dados Porta Lógica XOR

Base de dados:

X	Y	Classificação	
		Original	Perceptron
0	0	0	0
0	1	1	1
1	0	1	1
1	1	0	0

Perceptron Porta XOR		
Função de ativação: $f(W^T X) = \begin{cases} 1 & \text{se } \sum_{i=0}^n w_i x_i \geq \alpha \\ -1 & \text{caso contrário} \end{cases}$		
Dimensões	d	2
Taxa de aprendizagem	η	0.01
Coefficiente de saturação	α	0.5
Episódios/eras	E	20.000
Camadas e perceptrons	C	2 – [2, 1]
Pesos Sinápticos	W	(0.5577 0.5796 -0.0855)

Resultados:

		Classificação Correta		Acurácia
		0	1	
Perceptron	0	2	0	100%
	1	0	2	100%
Acurácia Total				100%

Os testes demonstraram que com 3 perceptrons, 1 saída e 1 camada oculta com 2 perceptrons podem representar uma porta lógica do tipo XOR.

Aluno Cristiano Lopes Moreira		RA/Matrícula 119103-0	Professor Dr Reinaldo Bianchi	Tipo Relatório de implementação	
Data 12/12/2019	Versão 1	Turma 2º. Semestre de 2019	Nome do arquivo PEL_208_Relatório_06_Cristiano_Moreira.doc		Página 10 (32)

4.2. Iris Fisher

Base de dados:

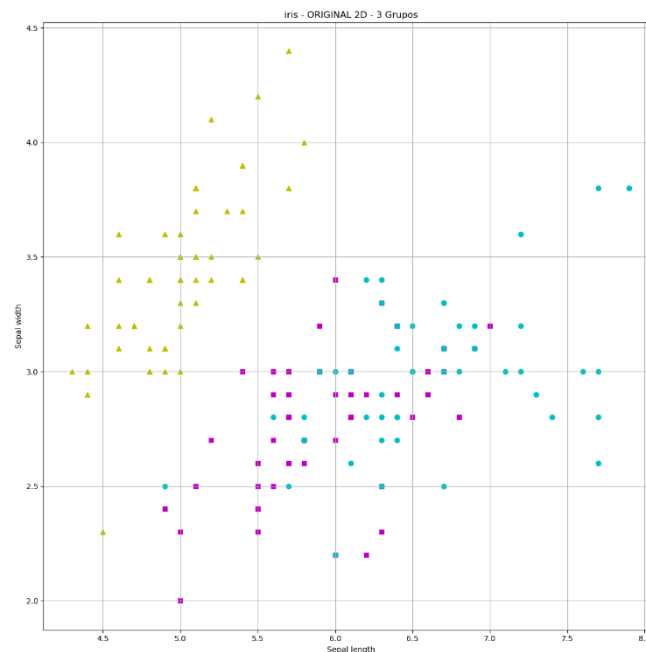


Fig4: Distribuição Original 3 dimensões visão em 2D da base de dados Iris-Fisher

O conjunto de dados contém 3 classes de 50 instâncias cada, em que cada classe se refere a um tipo de planta de íris, mostrados na Figura 2.

Uma classe é linearmente separável das outras duas e as demais não são linearmente separáveis.

Atributo previsto: classe da planta íris.

Informações da base:

- comprimento da sépala em cm
- largura da sépala em cm
- comprimento da pétala em cm
- largura da pétala em cm
- 3 classes: (Setosa, Versicolour e Virginica)

Aluno		RA/Matrícula		Professor	Tipo
Cristiano Lopes Moreira		119103-0		Dr Reinaldo Bianchi	Relatório de implementação
Data	Versão	Turma		Nome do arquivo	Página
12/12/2019	1	2º. Semestre de 2019		PEL_208_Relatório_06_Cristiano_Moreira.doc	11 (32)

MLP Base de Dados Iris

Função de ativação: $f(W^T X) = \begin{cases} 1 & \text{se } \sum_{i=0}^n w_i x_i \geq \alpha \\ -1 & \text{caso contrário} \end{cases}$

Dimensões	d	4
Taxa de aprendizagem	η	0.1
Coeficiente de saturação	α	0.5
Episódios/eras	E	20000
Camadas e perceptrons	C	2 – [8, 3]

Pesos Sinápticos

Camada 0 (oculta)

	w0	w1	w2	w3	w4
s1	-1.031	-0.777	-0.254	0.2736	-0.338
s2	1.1677	-3.927	1.8786	1.8931	-5.674
s3	-0.394	-1.273	-0.303	-0.788	-0.865
s4	0.0229	-2.113	-0.095	5.4231	-5.825
s5	-0.249	-0.61	-0.96	-0.657	-0.178
s6	-1.19	-0.964	-0.019	0.5845	-0.446
s7	0.1001	1.8105	-2.563	0.2854	-0.412
s8	-3.15	1.5346	4.0888	0.7443	-2.921

Camada 1 (Saída)

	w0	w1	w2	w3	w4	w5	w6	w7	w8
s1	-0.3128	-0.1943	-0.0146	-0.1404	1.0325	0.6996	0.1579	0.1473	-0.1173
s2	-0.7481	-0.2028	0.4779	-0.9876	-0.4585	0.1817	-0.9338	-0.4434	0.6369
s3	-0.8037	0.2650	-0.0339	1.1614	-0.4568	0.9623	-0.2987	0.4679	-0.6952

Resultados:

Para escolha da arquitetura da rede neural foram realizados 37 cenários alterando-se o número de camadas escondidas e o número de perceptrons dentro de cada camada, respeitando a regra de Hecht Nielsen com número de neurônios na camada escondida menor ou igual a 2 vezes a quantidade de variáveis acrescido de 1. Cada cenário foi executado com 2000 episódios/eras, com o resultado sendo a média de 10 repetições.

Pela metodologia apresentada o cenário com melhor resultado foi de 2 camadas, sendo 1 camada oculta com 8 perceptrons e uma camada de saída com 3 perceptrons ([8,3]). A essa arquitetura foram realizadas, para validação, 10 repetições dos 20.000 episódios para obter o cenário de melhor acerto dos itens da base de dados íris. A base foi dividida em tbTraining e tbValid, sendo a primeira 60 amostras aleatórias utilizadas para o treinamento, e a segunda a base para validação da classificação.

Aluno		RA/Matrícula		Professor		Tipo	
Cristiano Lopes Moreira		119103-0		Dr Reinaldo Bianchi		Relatório de implementação	
Data	Versão	Turma		Nome do arquivo			Página
12/12/2019	1	2º. Semestre de 2019		PEL_208_Relatório_06_Cristiano_Moreira.doc			12 (32)

Camada2	Camada1	Saída	Classificação Correta			Acurácia
			0	1	2	
0	0	3	49.8	09.2	46.2	70%
0	1	3	24.8	00.0	49.4	49%
0	2	3	38.6	04.6	48.5	61%
0	3	3	39.8	20.4	48.3	72%
0	4	3	49.1	13.9	48.2	74%
0	5	3	50.0	18.2	46.9	77%
0	6	3	44.4	27.9	46.3	79%
0	7	3	45.0	26.1	50.0	81%
0	8	3	42.5	30.8	50.0	82%
1	1	3	10.2	00.5	49.8	40%
1	2	3	24.6	00.7	50.0	50%
1	3	3	16.0	01.7	49.1	45%
1	4	3	21.2	12.4	49.7	56%
1	5	3	19.0	03.9	49.9	49%
1	6	3	06.0	05.1	48.6	40%
1	7	3	18.7	07.4	49.1	50%
2	1	3	10.0	00.9	49.1	40%
2	2	3	24.1	05.5	49.5	53%
2	3	3	14.9	15.1	49.3	53%
2	4	3	28.7	14.4	48.4	61%
2	5	3	29.8	17.4	48.9	64%
2	6	3	17.5	11.4	49.2	52%
3	1	3	24.7	07.3	49.8	55%
3	2	3	38.9	10.4	47.6	65%
3	3	3	34.1	13.6	49.1	65%
3	4	3	39.6	14.0	48.5	68%
3	5	3	33.1	23.0	47.9	69%
4	1	3	24.4	04.1	49.7	52%
4	2	3	34.9	04.5	49.4	59%
4	3	3	44.3	29.2	48.7	81%
4	4	3	39.7	18.9	47.5	71%
5	1	3	33.9	05.1	49.5	59%
5	2	3	47.3	08.4	47.6	69%
5	3	3	42.5	06.2	48.0	64%
6	1	3	31.4	09.4	49.0	60%
6	2	3	40.9	08.8	49.4	66%
7	1	3	44.7	07.5	47.7	67%

Aluno		RA/Matrícula		Professor		Tipo	
Cristiano Lopes Moreira		119103-0		Dr Reinaldo Bianchi		Relatório de implementação	
Data	Versão	Turma		Nome do arquivo		Página	
12/12/2019	1	2º. Semestre de 2019		PEL_208_Relatório_06_Cristiano_Moreira.doc		13 (32)	

Acurácia x Quantidade de pesos “w” na rede neural

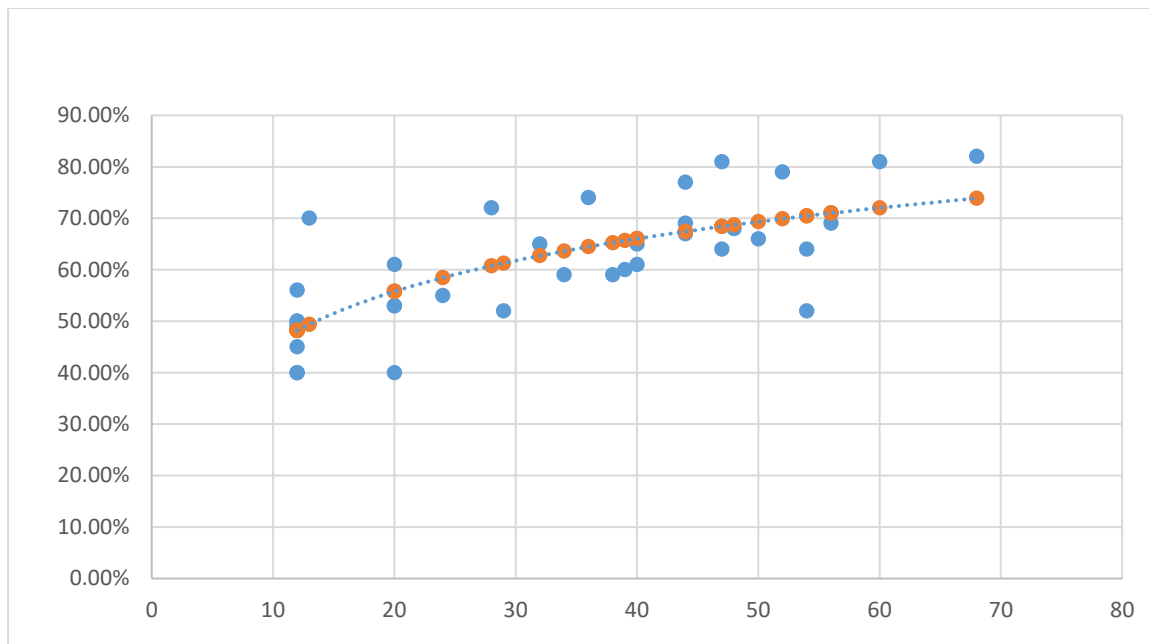


Fig5: Regressão logarítmica da interação de pesos e acurácia da RNA iris

Decomposição da acurácia em armonicadas por injeção de camadas

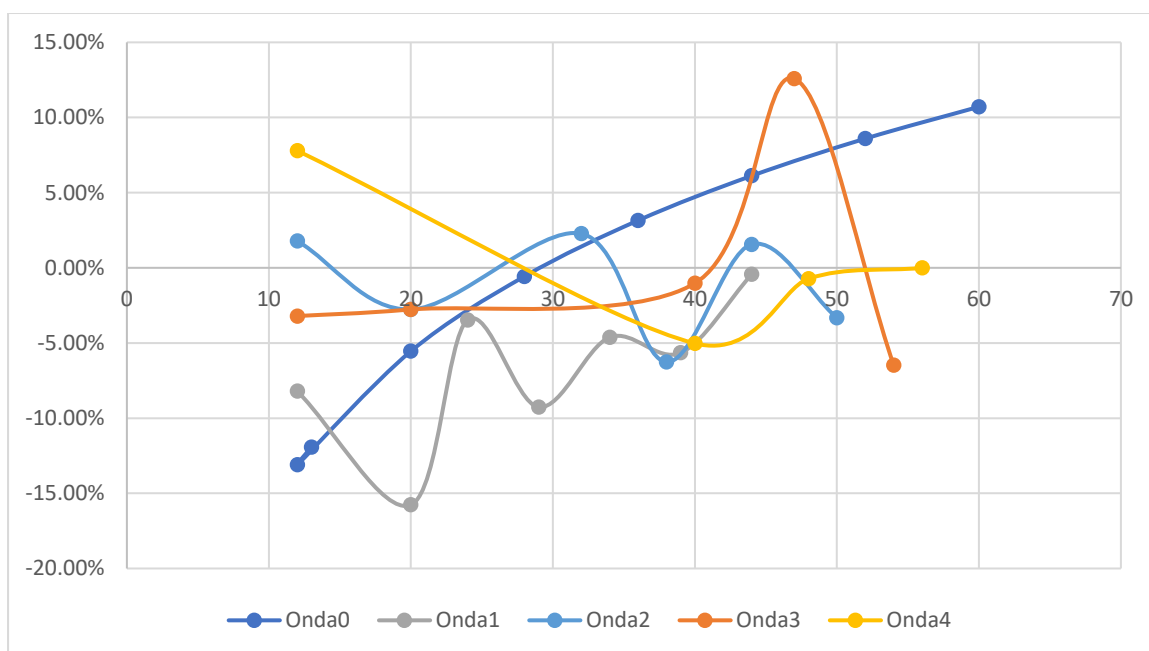


Fig6: Decomposição da curva de pesos x acurácia da RNA iris

Aluno		RA/Matrícula		Professor		Tipo	
Cristiano Lopes Moreira		119103-0		Dr Reinaldo Bianchi		Relatório de implementação	
Data	Versão	Turma		Nome do arquivo		Página	
12/12/2019	1	2º. Semestre de 2019		PEL_208_Relatório_06_Cristiano_Moreira.doc		14 (32)	

		Classificação			Acurácia
		0	1	2	
Perceptron	0	50	0	0	100%
	1	0	46	4	92%
	2	0	0	50	100%
Acurácia Total					97%

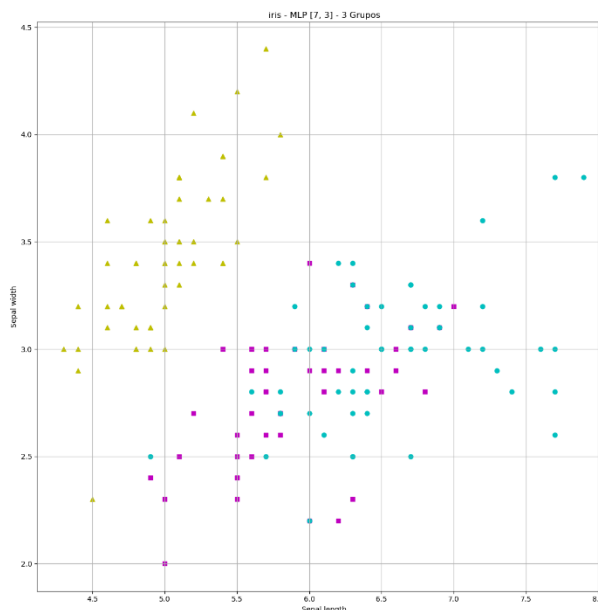


Fig7: MLP 3D visão em 2D da base de dados Iris

Pela tabela de confusão observar-se que o MLP é eficiente para classificação dos dados, apresentou 97% de acerto na classificação dos grupos da Iris: 0 Setosa, 1 Versicolour e 2 Virginica; em comparação com a classificação real informada pela base de dados.

Comprova-se também a regra de Hecht-Nielsen, que ter mais camadas neurais não é necessariamente o critério que irá gerar aumento significativo na acurácia. A arquitetura de maior acurácia foi de 1 camada oculta, e 7 neurônios, com resultados melhores que outras arquiteturas validadas com mais camadas e mais neurônios.

É possível observar, pelos gráficos da figura 5 e 6, que existe uma relação logarítmica entre a quantidade de pesos sinápticos e a acurácia máxima da RNA.

Aluno Cristiano Lopes Moreira		RA/Matrícula 119103-0	Professor Dr Reinaldo Bianchi	Tipo Relatório de implementação	
Data 12/12/2019	Versão 1	Turma 2º. Semestre de 2019	Nome do arquivo PEL_208_Relatório_06_Cristiano_Moreira.doc		Página 15 (32)

4.3. Seed

Base de dados:

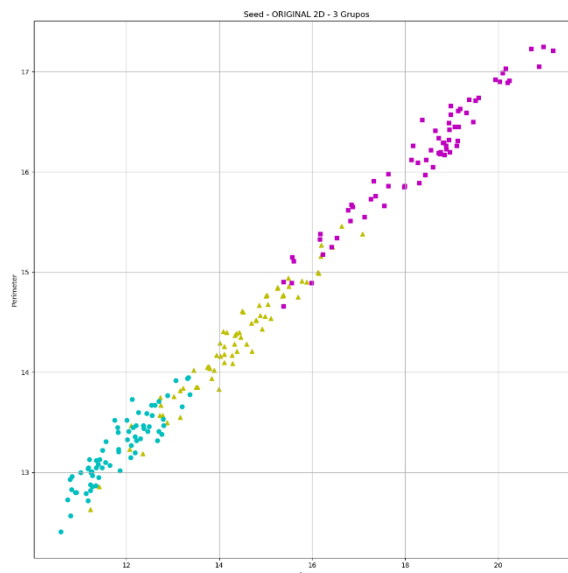


Fig8: Distribuição Original 6 dimensões visão em 2D da base de dados Seed

Base de dados do Instituto de Agrofísica da Academia Polonesa de Ciências de Lublin contendo amostras de grãos pertencentes a três variedades diferentes de trigo: Kama, Rosa e Canadense, com 70 elementos cada, selecionados aleatoriamente.

Informações dos atributos da base:

- Área
- Perímetro
- Compacidade $C = 4 \cdot \pi \cdot A / P^2$
- Comprimento do núcleo da semente
- Largura do núcleo da semente
- Coeficiente de assimetria
- Comprimento do sulco do núcleo da semente
- 3 classes: 1 para Kama, 2 para Rosa e 3 para Canadense

Aluno		RA/Matrícula		Professor		Tipo	
Cristiano Lopes Moreira		119103-0		Dr Reinaldo Bianchi		Relatório de implementação	
Data	Versão	Turma		Nome do arquivo			Página
12/12/2019	1	2º. Semestre de 2019		PEL_208_Relatório_06_Cristiano_Moreira.doc			16 (32)

MLP Base de Dados Seed														
Função de ativação: $f(W^T X) = \begin{cases} 1 & \text{se } \sum_{i=0}^n w_i x_i \geq \alpha \\ -1 & \text{caso contrário} \end{cases}$														
Dimensões	d	7												
Taxa de aprendizagem	η	0.01												
Coeficiente de saturação	α	0.5												
Episódios/eras	E	20000												
Camadas e perceptrons	C	2 – [12, 3]												
Pesos Sinápticos														
Camada 0 (oculta)														
		w0	w1	w2	w3	w4	w5	w6	w7					
s1		-1.2364	0.9222	0.9751	1.5241	1.2782	-0.7813	-0.7917	-0.2313					
s2		0.7259	-0.3115	-0.3147	0.697	0.4655	-0.0312	0.1364	-0.9728					
s3		-0.8776	-0.4775	-0.8117	-0.082	0.583	-0.5861	0.2026	0.6181					
s4		-0.9401	-0.3199	-0.1001	0.9772	0.5294	0.3886	-0.6248	-0.2121					
s5		-0.8973	-0.8446	0.3905	-0.0856	0.0104	0.9989	0.5088	-0.6116					
s6		-0.8347	-0.2307	-0.9532	0.1822	-0.5769	0.6305	0.0042	0.4073					
s7		0.6941	0.9166	0.9602	-0.8359	0.9166	0.5949	-0.5811	-0.235					
s8		-0.8932	0.145	0.8277	-0.3517	0.752	-0.8844	-0.8983	-0.2379					
s9		-0.7268	-0.2619	-0.9109	0.1781	0.8256	-0.2996	-0.0176	-0.6575					
s10		2.3999	-2.3851	0.6133	0.1904	-0.9176	1.8601	-1.7094	0.5029					
s11		0.2517	0.3121	0.7417	0.7759	-0.0851	0.3027	-0.5393	-0.8723					
s12		-0.9691	0.1802	-0.1828	0.8398	0.5261	1.164	-0.7012	0.4764					
Camada 1 (Saída)														
		w0	w1	w2	w3	w4	w5	w6	w7	w8	w9	w10	w11	w12
s1		0.0639	-0.9081	0.9605	-0.7115	0.8848	-0.9055	0.0944	-0.5805	0.3691	0.041	0.8168	-0.0485	-0.0592
s2		-0.162	0.0294	-0.9134	0.9862	0.7727	-0.9031	-0.4355	-0.9127	-0.1907	-0.0801	0.2481	-0.0639	0.2913
s3		0.0758	-0.4038	0.9839	-0.0967	-0.467	-0.5882	0.9589	0.2473	0.5242	-0.0029	-0.305	0.1411	-0.3251

Resultados:

Para escolha da arquitetura da rede neural foram realizados 106 cenários alterando-se o número de camadas escondidas e o número de perceptrons dentro de cada camada, respeitando a regra de Hecht Nielsen com número de neurônios na camada escondida menor ou igual a 2 vezes a quantidade de variáveis acrescido de 1. Cada cenário foi executado com 2000 episódios/eras.

Pela metodologia apresentada o cenário com melhor resultado foi de 2 camadas, sendo 1 camada oculta com 12 perceptrons e uma camada de saída com 3 perceptrons ([12,3]). A essa arquitetura foram realizadas 10 repetições dos 20.000 episódios para obter o cenário de melhor acerto dos itens da base de dados seed. A base foi dividida em tbTraining e tbValid, sendo a primeira 60% das amostras aleatórias utilizadas para o treinamento, e a segunda a base para validação.

Aluno		RA/Matrícula		Professor		Tipo	
Cristiano Lopes Moreira		119103-0		Dr Reinaldo Bianchi		Relatório de implementação	
Data	Versão	Turma		Nome do arquivo		Página	
12/12/2019	1	2º. Semestre de 2019		PEL_208_Relatório_06_Cristiano_Moreira.doc		17 (32)	

			Classificação Correta			Acurácia
Camada2	Camada1	Saida	0	1	2	
0	0	3	38.2	54.2	52.4	69%
0	1	3	2.2	69.2	23.6	45%
0	2	3	0	70	0	33%
0	3	3	15	68.4	1.6	40%
0	4	3	12	65.8	36.2	54%
0	5	3	6	64.6	15	41%
0	6	3	14.2	63.6	39	56%
0	7	3	40.4	56	47	68%
0	8	3	41	60.4	51.8	73%
0	9	3	40.2	66.6	13	57%
0	10	3	34	65.6	48.6	71%
0	11	3	33	60.6	51.8	69%
0	12	3	41.8	68.2	61.2	82%
0	13	3	29	63.8	50.4	68%
0	14	3	37.2	66	49.8	73%
1	1	3	0	70	0.4	34%
1	2	3	5.2	65.6	14.8	41%
1	3	3	1	70	0	34%
1	4	3	0.2	67	36.6	49%
1	5	3	2.4	67.4	10.8	38%
1	6	3	13.8	68.2	12.6	45%
1	7	3	1.2	70	0	34%
1	8	3	4	68	0	34%
1	9	3	0	70	12.4	39%
1	10	3	0	67.6	14.2	39%
1	11	3	3	63.6	0	32%
1	12	3	13	65.8	49.2	61%
1	13	3	0	69.2	19	42%
2	1	3	5	69.8	0	36%
2	2	3	1.4	65.8	17.2	40%
2	3	3	17.8	68.6	3.4	43%
2	4	3	0	68.2	25.2	44%
2	5	3	22.8	58.8	40.2	58%
2	6	3	9	66	26.6	48%
2	7	3	16.4	64	38.8	57%
2	8	3	10.4	68.8	14	44%

Aluno		RA/Matrícula		Professor		Tipo	
Cristiano Lopes Moreira		119103-0		Dr Reinaldo Bianchi		Relatório de implementação	
Data	Versão	Turma		Nome do arquivo		Página	
12/12/2019	1	2º. Semestre de 2019		PEL_208_Relatório_06_Cristiano_Moreira.doc		18 (32)	

			Classificação Correta			Acurácia
Camada2	Camada1	Saida	0	1	2	
2	9	3	23	67.6	19	52%
2	10	3	3.8	68.4	22.8	45%
2	11	3	16.6	63.6	26.8	51%
2	12	3	2.2	70	0	34%
3	1	3	0.2	69	37.8	51%
3	2	3	17.8	69.4	1.4	42%
3	3	3	0	68	26.4	45%
3	4	3	0	60.6	37.4	47%
3	5	3	16.8	62	27	50%
3	6	3	19.2	66.4	33.6	57%
3	7	3	6.2	61.2	11	37%
3	8	3	13.4	62.2	15.8	44%
3	9	3	11	53.2	61.8	60%
3	10	3	18.8	65	27.2	53%
3	11	3	10.6	69.6	24	50%
4	1	3	15.6	65.2	10	43%
4	2	3	9.2	69	36.4	55%
4	3	3	0.4	70	24.6	45%
4	4	3	10.6	69	26.2	50%
4	5	3	10.4	66.8	53	62%
4	6	3	15.8	67.2	48.2	62%
4	7	3	10.6	59.8	15.4	41%
4	8	3	5.8	69.6	44.4	57%
4	9	3	11.4	63.2	27.2	48%
4	10	3	16.4	68.2	14	47%
5	1	3	20.4	66.6	19.6	51%
5	2	3	25.8	65.8	29.2	58%
5	3	3	18.4	56	51.8	60%
5	4	3	12	65.8	28.4	51%
5	5	3	14.8	64.2	52.2	62%
5	6	3	3.8	62.6	36	49%
5	7	3	38	68	41.4	70%
5	8	3	38.8	62.2	61.4	77%
5	9	3	19.6	65.8	52.2	66%
6	1	3	0	63.4	45.4	52%
6	2	3	12.6	58.8	40	53%

Aluno		RA/Matrícula		Professor		Tipo	
Cristiano Lopes Moreira		119103-0		Dr Reinaldo Bianchi		Relatório de implementação	
Data	Versão	Turma		Nome do arquivo		Página	
12/12/2019	1	2º. Semestre de 2019		PEL_208_Relatório_06_Cristiano_Moreira.doc		19 (32)	

			Classificação Correta			Acurácia
Camada2	Camada1	Saida	0	1	2	
6	3	3	0	65.6	26	44%
6	4	3	14	66.4	32.8	54%
6	5	3	37.6	65.8	14	56%
6	6	3	16.4	63	43.2	58%
6	7	3	26.6	59.4	58.6	69%
6	8	3	1	69.6	38.2	52%
7	1	3	2.2	64	25.2	44%
7	2	3	10.4	63.2	36.4	52%
7	3	3	0.4	65.4	28.4	45%
7	4	3	9.4	68.2	38	55%
7	5	3	13.8	66	46.4	60%
7	6	3	37	67.2	47.8	72%
7	7	3	23.6	67.6	50.6	68%
8	1	3	4	70	25.6	47%
8	2	3	17.6	62.2	47.6	61%
8	3	3	20	64.2	60	69%
8	4	3	18.8	68.4	46.4	64%
8	5	3	32.8	59.4	49.4	67%
8	6	3	22.2	64.2	54.6	67%
9	1	3	8.6	64	27	47%
9	2	3	12.8	59	51.6	59%
9	3	3	26.4	62	53	67%
9	4	3	19	63	40.8	58%
9	5	3	25.6	64.8	48.8	66%
10	1	3	0	69.2	38	51%
10	2	3	33.8	61.2	40.6	65%
10	3	3	31.6	69.2	43.8	69%
10	4	3	23.2	60.8	56	67%
11	1	3	2.2	63.6	26.4	44%
11	2	3	25.8	66.6	41.2	64%
11	3	3	16.2	68.4	35	57%
12	1	3	1	64.4	47.6	54%
12	2	3	10.2	56.4	38.2	50%
13	1	3	4.4	65.4	1.6	34%

Aluno		RA/Matrícula		Professor		Tipo	
Cristiano Lopes Moreira		119103-0		Dr Reinaldo Bianchi		Relatório de implementação	
Data	Versão	Turma		Nome do arquivo		Página	
12/12/2019	1	2º. Semestre de 2019		PEL_208_Relatório_06_Cristiano_Moreira.doc		20 (32)	

Acurácia x Quantidade de pesos “w” na rede neural

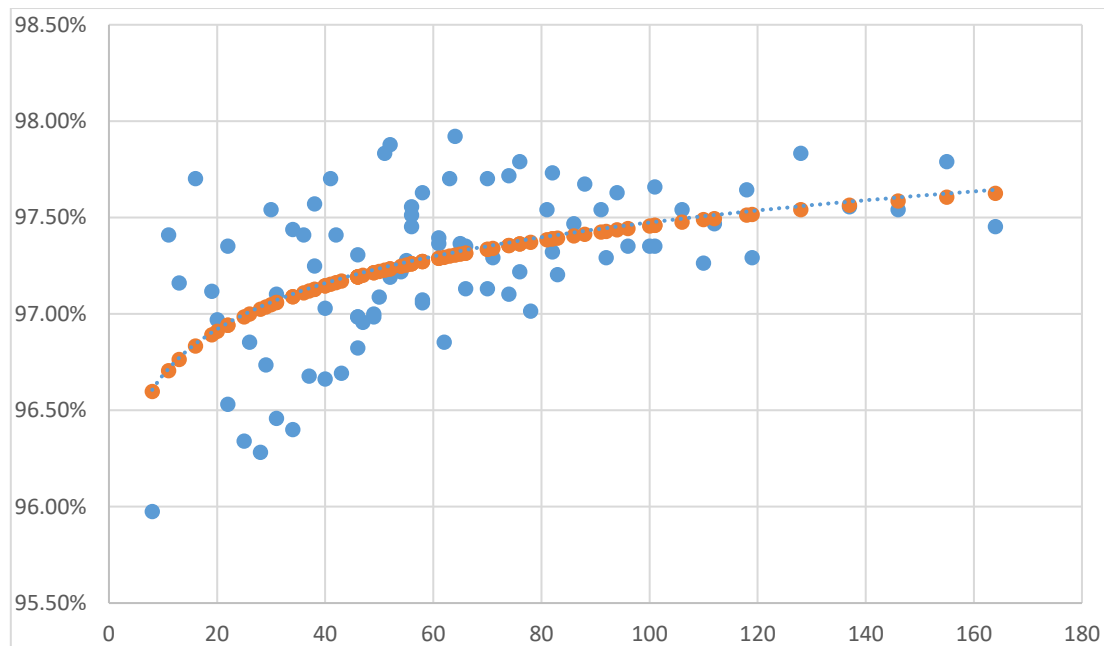


Fig9: Regressão logarítmica da interação de pesos e acurácia da RNA Seed

Decomposição da acurácia em armonicis por injeção de camadas

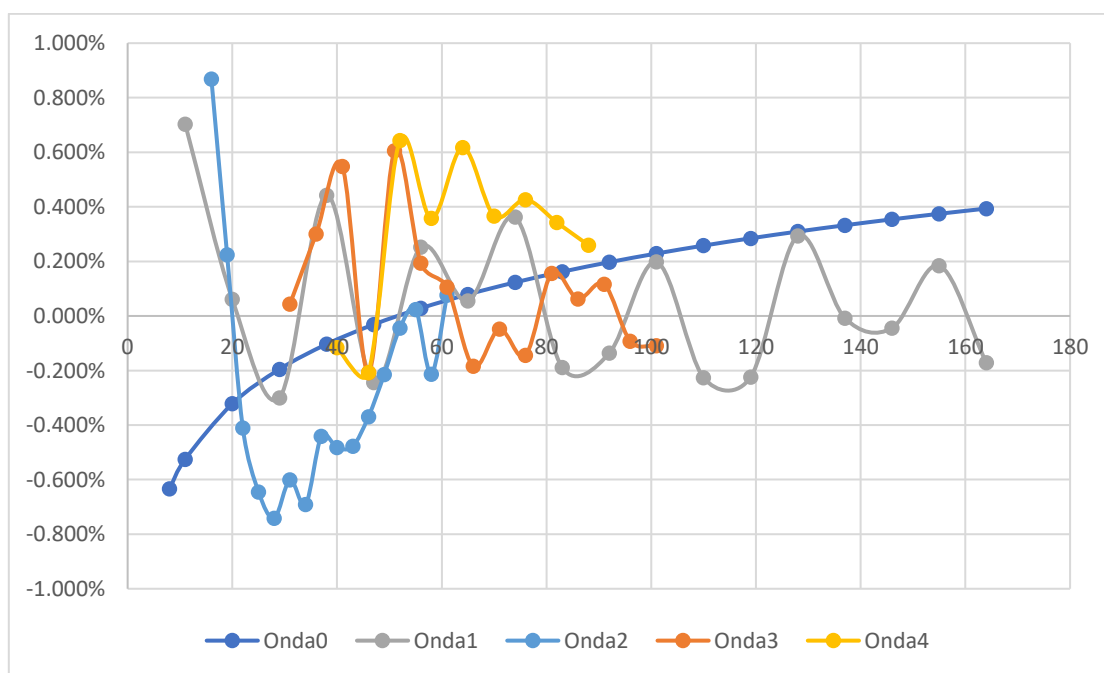


Fig10: Decomposição da curva de pesos x acurácia da RNA Seed

Aluno Cristiano Lopes Moreira		RA/Matrícula 119103-0	Professor Dr Reinaldo Bianchi	Tipo Relatório de implementação	
Data 12/12/2019	Versão 1	Turma 2º. Semestre de 2019	Nome do arquivo PEL_208_Relatório_06_Cristiano_Moreira.doc		Página 21 (32)

		Classificação			Acurácia
		0	1	2	
Perceptron	0	62	1	1	88.57%
	1	3	65	0	92.86%
	2	2	10	58	82.86%
Acurácia Total					88.10%

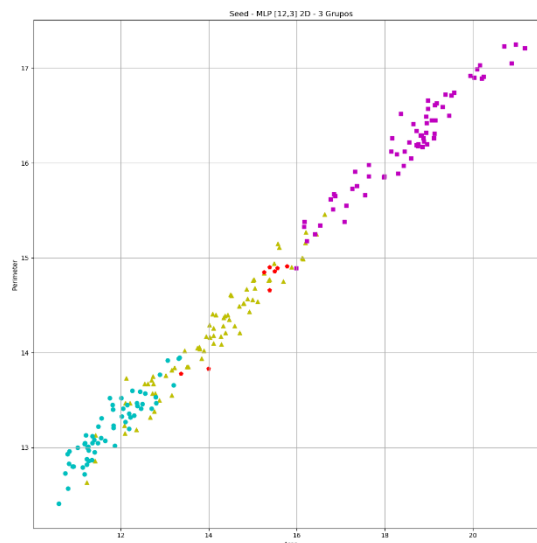


Fig11: MLP 9D visão em 2D da base de dados Seed

Pela tabela de confusão observar-se que o MLP é eficiente para classificação dos dados, apresentou 88.1% de acerto na classificação entre os grãos de trigo do tipo1 Kama, 2 para Rosa e 3 para Canadense; em comparação com a classificação real informada pela base de dados.

Comprova-se também a regra de Hecht-Nielsen, que ter mais camadas neurais não é necessariamente o critério que irá gerar maior acurácia. A arquitetura de maior acurácia foi de 1 camada oculta, e 12 neurônios, com resultados melhores que outras arquiteturas validadas com mais camadas e mais neurônios.

Pela decomposição da acurácia x quantidade de pesos sinápticos w em harmônicas, figuras 9 e 10, observa-se um padrão logaritmo e, sobre este padrão logaritmo, uma variação senoidal com decréscimo de acurácia para os casos de quantidade positiva de neurônios nas camadas ocultas. Observa-se também que acima de 120 pesos sinápticos a acurácia é regida majoritariamente pela primeira camada oculta (Onda0 e Onda1)

Aluno		RA/Matrícula		Professor		Tipo	
Cristiano Lopes Moreira		119103-0		Dr Reinaldo Bianchi		Relatório de implementação	
Data	Versão	Turma		Nome do arquivo		Página	
12/12/2019	1	2º. Semestre de 2019		PEL_208_Relatório_06_Cristiano_Moreira.doc		22 (32)	

4.4. Breast Cancer Wisconsin (Original) Data Set

Base de dados:

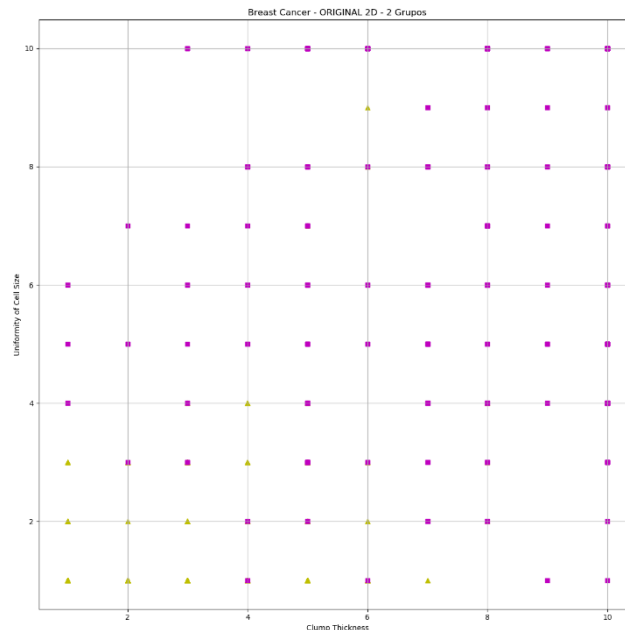


Fig12: Distribuição Original 9 dimensões visão em 2D da base de dados Breast Cancer Wisconsin

A base de dados das amostras do Dr. Wolberg, W.H. & Mangasarian, O.L. (1990) para criação de um método multi-superfície de separação de padrões para diagnóstico médico aplicado à citologia mamária; contém 699 casos clínicos de suspeitas de câncer de mama, sendo 15 sem a informação de núcleos nus (bare nuclei), e por isto removidas desta análise.

Informações dos atributos da base:

- Espessura do grupo
- Uniformidade do tamanho da célula
- Uniformidade da forma da célula
- Adesão marginal
- Tamanho único de célula epitelial
- Núcleos Nus
- Cromatina Branda
- Núcleos normais
- Mitoses
- 2 classes: 1 para benignos, 2 para malignos

Aluno		RA/Matrícula		Professor		Tipo	
Cristiano Lopes Moreira		119103-0		Dr Reinaldo Bianchi		Relatório de implementação	
Data	Versão	Turma		Nome do arquivo			Página
12/12/2019	1	2º. Semestre de 2019		PEL_208_Relatório_06_Cristiano_Moreira.doc			23 (32)

MLP Base de Dados Breast-Cancer-Wisconsin

$$\text{Função de ativação: } f(W^T X) = \begin{cases} 1 & \text{se } \sum_{i=0}^n w_i x_i \geq \alpha \\ -1 & \text{caso contrário} \end{cases}$$

Dimensões	d	9
Taxa de aprendizagem	η	0.1
Coeficiente de saturação	α	0.5
Episódios/eras	E	20000
Camadas e perceptrons	C	2 – [7, 7, 3]

Pesos Sinápticos

Camada 2 (oculta)

	w0	w1	w2	w3	w4	w5	w6	w7	w8	w9
s1	0.5826	2.0981	0.898	0.502	-1.2732	1.3147	-2.5564	1.2216	1.7634	-4.2543
s2	-1.7447	2.2264	-0.0934	-0.8276	1.1359	2.3703	0.7592	0.2083	1.0817	-3.0484
s3	1.3116	-0.4157	-0.3501	1.3079	0.0157	0.1409	0.3807	0.1364	0.3166	-6.1047
s4	2.0556	3.3217	1.8641	0.4558	-2.8666	0.504	2.0026	-3.7872	2.557	-7.7306
s5	-0.2064	0.523	0.2522	-0.1758	0.7546	1.2555	0.123	0.2905	-0.0582	-4.4795
s6	2.5837	2.1551	0.9225	-0.6503	-0.9054	-3.7182	6.4885	1.0579	-3.302	0.2999
s7	0.3415	-1.651	0.5354	0.0672	0.369	0.0963	2.06	1.2628	1.5129	-5.5938

Camada 1 (oculta)

	w0	w1	w2	w3	w4	w5	w6	w7
s1	-0.3543	-0.7807	-0.8545	-0.241	-0.7215	-0.4166	-0.688	-0.3443
s2	-0.4269	-0.7416	-0.8551	-0.2584	-0.617	-0.4244	-0.7818	-0.3809
s3	-0.3672	-0.6319	-0.8523	-0.2417	-0.6924	-0.5683	-0.6845	-0.4384
s4	-0.5012	-0.6979	-0.8337	-0.3491	-0.6481	-0.5348	-0.6433	-0.3811
s5	-0.5062	-0.6219	-0.8333	-0.416	-0.6845	-0.4323	-0.739	-0.3929
s6	-0.3783	-0.6628	-0.7698	-0.3471	-0.7343	-0.5602	-0.7817	-0.3103
s7	-0.3596	-0.6788	-0.9506	-0.2719	-0.656	-0.4082	-0.6747	-0.4441

Camada 0 (Saída)

	w0	w1	w2	w3	w4	w5	w6	w7
s1	-0.4829	-0.317	-0.2246	-0.3691	-0.3677	-0.4408	-0.173	0.0398

Resultados:

Para escolha da arquitetura da rede neural foram realizados 158 cenários alterando-se o número de camadas escondidas e o número de perceptrons dentro de cada camada, respeitando a regra de Hecht Nielsen com número de neurônios na camada escondida menor ou igual a 2 vezes a quantidade de variáveis acrescido de 1. Cada cenário foi executado com 2000 episódios/eras.

Pela metodologia apresentada o cenário com melhor resultado foi o [11,7,1], porém, pela diferença de ganho da acurácia ser muito pequena, 0,06%, e o custo em tempo em tempo superior, foi escolhido [7 ,7 1] sendo 2 camada oculta com 7 perceptrons e uma cada de saída com 1 perceptrons. A essa arquitetura foram

Aluno		RA/Matrícula		Professor		Tipo	
Cristiano Lopes Moreira		119103-0		Dr Reinaldo Bianchi		Relatório de implementação	
Data	Versão	Turma		Nome do arquivo		Página	
12/12/2019	1	2º. Semestre de 2019		PEL_208_Relatório_06_Cristiano_Moreira.doc		24 (32)	

realizadas, para validação, 10 repetições dos 20.000 episódios para obter o cenário de melhor acerto dos itens da base de dados breast_cancer. A base foi dividida em tbTraining e tbValid, sendo a primeira 50% das amostras aleatórias utilizadas para o treinamento, e a segunda a base para validação.

Camada1	Camada2	Saída	Classificação Correta		Precisão
			0	1	
0	0	1	437.1	218.4	95.97%
0	1	1	431.6	233.7	97.41%
0	2	1	433.9	228.4	96.97%
0	3	1	435.3	225.4	96.73%
0	4	1	433.8	232.6	97.57%
0	5	1	434.7	227.5	96.95%
0	6	1	431.1	234.9	97.51%
0	7	1	432.2	232.8	97.36%
0	8	1	434.2	233.2	97.72%
0	9	1	435.1	228.8	97.20%
0	10	1	432	232.5	97.29%
0	11	1	431.4	235.6	97.66%
0	12	1	436	228.3	97.26%
0	13	1	434.1	230.4	97.29%
0	14	1	435.2	233	97.83%
0	15	1	433.9	232.4	97.55%
0	16	1	434.4	231.8	97.54%
0	17	1	434.3	233.6	97.79%
0	18	1	435.3	230.3	97.45%
1	1	1	431.2	232.4	97.16%
1	2	1	430.9	236.4	97.70%
1	3	1	433.2	230.1	97.12%
1	4	1	429.4	229.9	96.53%
1	5	1	433.6	224.4	96.34%
1	6	1	435.2	222.4	96.28%
1	7	1	431.4	227.4	96.46%
1	8	1	434.2	224.2	96.40%
1	9	1	433.7	226.6	96.68%
1	10	1	434	226.2	96.66%
1	11	1	433.5	226.9	96.69%

Aluno		RA/Matrícula		Professor		Tipo	
Cristiano Lopes Moreira		119103-0		Dr Reinaldo Bianchi		Relatório de implementação	
Data	Versão	Turma		Nome do arquivo		Página	
12/12/2019	1	2º. Semestre de 2019		PEL_208_Relatório_06_Cristiano_Moreira.doc		25 (32)	

			Classificação Correta		
Camada1	Camada2	Saida	0	1	Precisão
1	12	1	433.5	227.8	96.82%
1	13	1	432.9	229.6	97.00%
1	14	1	432.7	231.1	97.19%
1	15	1	432.8	231.6	97.28%
1	16	1	432.9	230	97.06%
1	17	1	432.5	232.5	97.36%
2	1	1	432.9	232	97.35%
2	2	1	434.2	227.3	96.85%
2	3	1	432.4	233.8	97.54%
2	4	1	429.9	235.6	97.44%
2	5	1	429.5	234.7	97.25%
2	6	1	428.3	237	97.41%
2	7	1	428.4	236.2	97.31%
2	8	1	432.2	230.9	97.09%
2	9	1	431.1	232.9	97.22%
2	10	1	432.6	230.4	97.07%
2	11	1	429.4	232.1	96.85%
2	12	1	430.9	234	97.35%
2	13	1	433.3	230.1	97.13%
2	14	1	433.8	229.4	97.10%
2	15	1	432	230.6	97.01%
2	16	1	429.8	234.9	97.32%
3	1	1	430.9	232.3	97.10%
3	2	1	430.1	235.2	97.41%
3	3	1	432.6	234.7	97.70%
3	4	1	434.7	227.7	96.98%
3	5	1	432	236.2	97.83%
3	6	1	433.4	232.2	97.45%
3	7	1	432.5	232.7	97.39%
3	8	1	432.9	230.5	97.13%
3	9	1	432.4	232.1	97.29%
3	10	1	431.6	232.4	97.22%
3	11	1	433.2	233	97.54%
3	12	1	430.2	235.5	97.47%
3	13	1	434.8	231.4	97.54%
3	14	1	433.8	231.1	97.35%

Aluno		RA/Matrícula		Professor		Tipo	
Cristiano Lopes Moreira		119103-0		Dr Reinaldo Bianchi		Relatório de implementação	
Data	Versão	Turma		Nome do arquivo		Página	
12/12/2019	1	2º. Semestre de 2019		PEL_208_Relatório_06_Cristiano_Moreira.doc		26 (32)	

			Classificação Correta		
Camada1	Camada2	Saida	0	1	Precisão
3	15	1	432.5	232.4	97.35%
4	1	1	434.3	228.4	97.03%
4	2	1	427.7	234.7	96.98%
4	3	1	432.3	236.2	97.88%
4	4	1	431.9	234.9	97.63%
4	5	1	431.2	237.6	97.92%
4	6	1	433.1	234.2	97.70%
4	7	1	435.9	232	97.79%
4	8	1	430.8	236.7	97.73%
4	9	1	429.4	237.7	97.67%
4	10	1	434.2	232.6	97.63%
4	11	1	436.4	228.5	97.35%
4	12	1	434.9	231.3	97.54%
4	13	1	434.2	231.5	97.47%
4	14	1	433.6	233.3	97.64%
5	1	1	431.5	230.9	96.98%
5	2	1	430.8	235.5	97.55%
5	3	1	431.3	236	97.70%
5	4	1	432.8	235.7	97.88%
5	5	1	433.2	235.7	97.94%
5	6	1	431.3	235.8	97.67%
5	7	1	433.5	234.6	97.82%
5	8	1	430.6	234.8	97.42%
5	9	1	437.1	228.8	97.50%
5	10	1	434.2	231.8	97.51%
5	11	1	435.9	231.2	97.67%
5	12	1	437.8	224.7	97.00%
5	13	1	439.1	224.9	97.22%
6	1	1	430.1	234	97.23%
6	2	1	431.8	235.4	97.69%
6	3	1	429.9	235.5	97.42%
6	4	1	430	234.6	97.31%
6	5	1	434.1	232	97.53%
6	6	1	435.2	232.3	97.73%
6	7	1	434.6	229.9	97.29%
6	8	1	435.2	232.1	97.70%

Aluno		RA/Matrícula		Professor		Tipo	
Cristiano Lopes Moreira		119103-0		Dr Reinaldo Bianchi		Relatório de implementação	
Data	Versão	Turma		Nome do arquivo		Página	
12/12/2019	1	2º. Semestre de 2019		PEL_208_Relatório_06_Cristiano_Moreira.doc		27 (32)	

			Classificação Correta		
Camada1	Camada2	Saida	0	1	Precisão
6	9	1	432.9	235.7	97.89%
6	10	1	436.7	229.1	97.48%
6	11	1	435.4	233	97.86%
6	12	1	435.6	230	97.45%
7	1	1	433.7	228.6	96.97%
7	2	1	426.9	236.8	97.17%
7	3	1	430.5	234.4	97.35%
7	4	1	431.7	235.4	97.67%
7	5	1	429.8	235.1	97.35%
7	6	1	433.9	235	97.94%
7	7	1	436.4	232.7	97.96%
7	8	1	432.2	236.2	97.86%
7	9	1	435.4	231.8	97.69%
7	10	1	435.8	229.5	97.41%
7	11	1	434.9	233.1	97.80%
8	1	1	428.7	232.7	96.84%
8	2	1	433.3	233.3	97.60%
8	3	1	432.1	233.8	97.50%
8	4	1	432	234.1	97.53%
8	5	1	435.4	231.9	97.70%
8	6	1	435.3	230.8	97.53%
8	7	1	434.5	233.8	97.85%
8	8	1	434.9	231	97.50%
8	9	1	432.6	233	97.45%
8	10	1	436.2	231.1	97.70%
9	1	1	430	233.3	97.12%
9	2	1	429.1	234.8	97.20%
9	3	1	431.6	231.3	97.06%
9	4	1	433.1	234.3	97.72%
9	5	1	431.3	237	97.85%
9	6	1	433.8	234.4	97.83%
9	7	1	435.8	231.6	97.72%
9	8	1	436.5	230.9	97.72%
9	9	1	434.7	232.8	97.73%
10	1	1	431.2	233.3	97.29%
10	2	1	432.3	233.7	97.51%

Aluno		RA/Matrícula		Professor		Tipo	
Cristiano Lopes Moreira		119103-0		Dr Reinaldo Bianchi		Relatório de implementação	
Data	Versão	Turma		Nome do arquivo		Página	
12/12/2019	1	2º. Semestre de 2019		PEL_208_Relatório_06_Cristiano_Moreira.doc		28 (32)	

			Classificação Correta		
Camada1	Camada2	Saida	0	1	Precisão
10	3	1	434	232	97.51%
10	4	1	434.9	232.7	97.75%
10	5	1	433.7	233.4	97.67%
10	6	1	432.8	233	97.48%
10	7	1	434.3	233.7	97.80%
10	8	1	433	236.3	97.99%
11	1	1	429.5	233.1	97.01%
11	2	1	431.7	234.3	97.51%
11	3	1	433.1	235.4	97.88%
11	4	1	436.1	234.2	98.14%
11	5	1	431.9	235.5	97.72%
11	6	1	433.6	233.9	97.73%
11	7	1	435.5	234	98.02%
12	1	1	431	234	97.36%
12	2	1	428.4	236.4	97.34%
12	3	1	431.5	236.6	97.82%
12	4	1	433.8	234	97.77%
12	5	1	434.7	233.7	97.86%
12	6	1	434.8	233.5	97.85%

Aluno		RA/Matrícula		Professor		Tipo	
Cristiano Lopes Moreira		119103-0		Dr Reinaldo Bianchi		Relatório de implementação	
Data	Versão	Turma		Nome do arquivo		Página	
12/12/2019	1	2º. Semestre de 2019		PEL_208_Relatório_06_Cristiano_Moreira.doc		29 (32)	

Acurácia x Quantidade de pesos “w” na rede neural

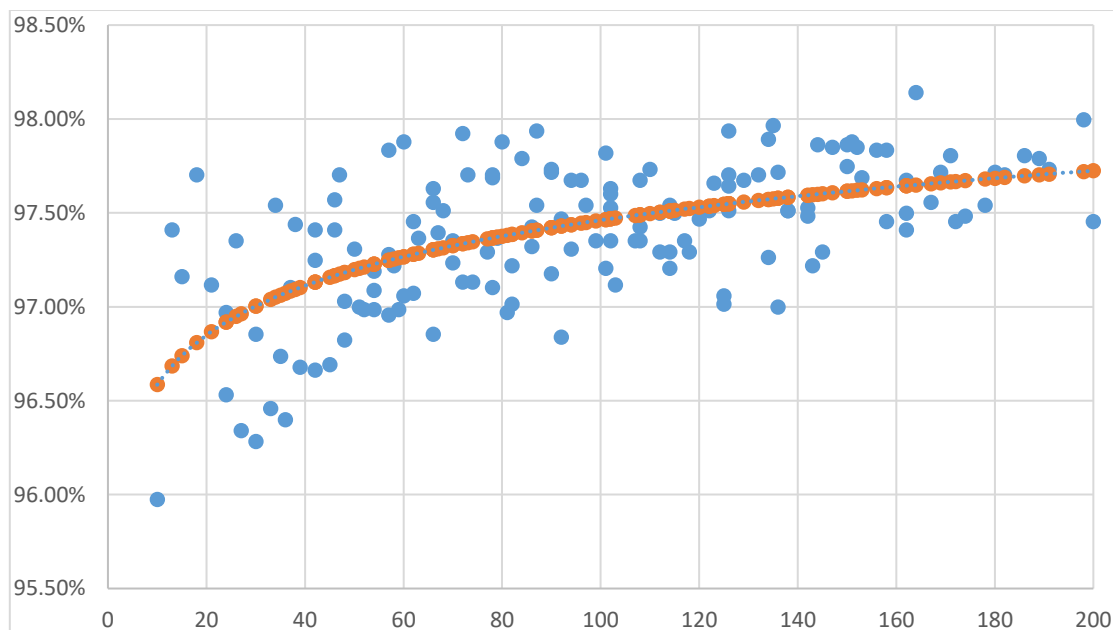


Fig13: Regressão logarítmica da interação de pesos e acurácia da RNA Breast-Cancer-Wisconsin

Decomposição da acurácia em armonicas por injeção de camadas

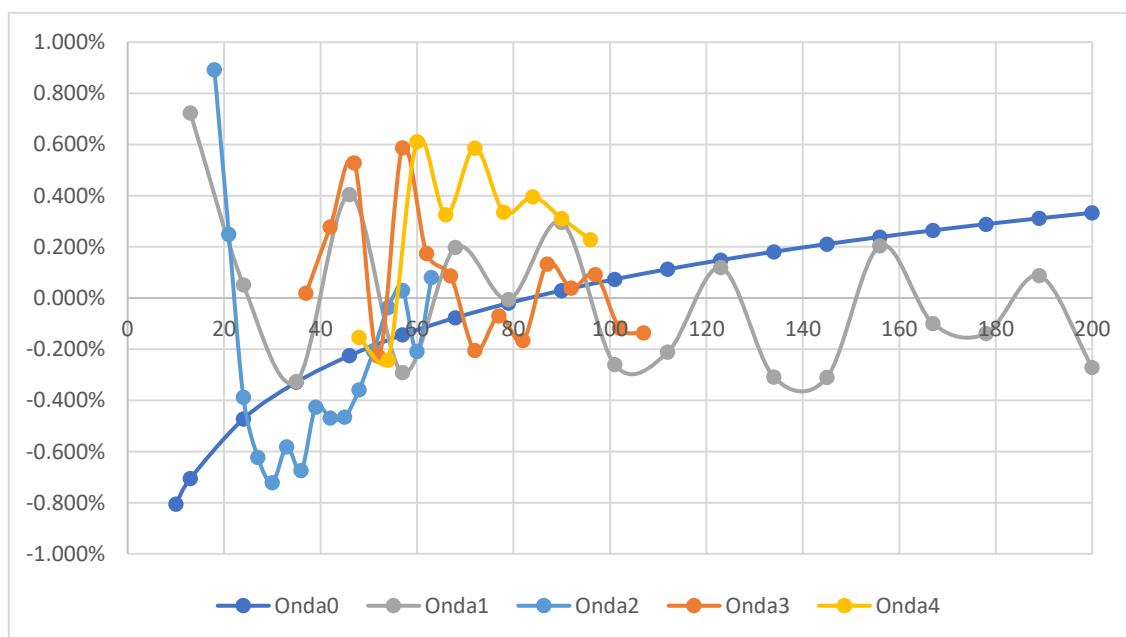


Fig14: Decomposição da curva de pesos x acurácia da RNA Breast-Cancer-Wisconsin

Aluno		RA/Matrícula	Professor	Tipo	
Cristiano Lopes Moreira		119103-0	Dr Reinaldo Bianchi	Relatório de implementação	
Data	Versão	Turma	Nome do arquivo		Página
12/12/2019	1	2º. Semestre de 2019	PEL_208_Relatório_06_Cristiano_Moreira.doc		30 (32)

		Classificação Correta		Acurácia
		0	1	
Perceptron	0	436	2	98.20%
	1	8	237	99.16%
Acurácia Total				98.54%

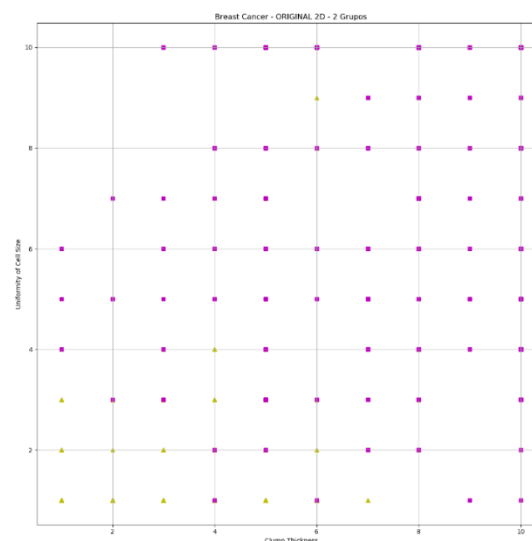


Fig13: MLP 9D visão em 2D da db Breast Cancer

Pela tabela de confusão observar-se que o MLP é eficiente para classificação dos dados com acurácia geral da identificação entre os dois grupos pelo modelo da RNA MLP foi de 98.54%, superior ao índice alvo de 95% (alvo de experimentos médicos), sendo maior nos acertos do grupo 1, tumor maligno, com 99.16% de acerto, dentro das expectativas para possibilitar diagnósticos precisos e rápidos.

Comprova-se também a regra de Hecht-Nielsen, que ter mais camadas neurais não é necessariamente o critério que irá gerar maior acurácia significativa. A arquitetura com muitas camadas e muitos acréscimos de perceptrons apresentaram acréscimos poucos significantes na ordem de 0.06%, 0.03%, [11,7,1] e [10.8,1], respectivamente.

Pela decomposição da acurácia x quantidade de pesos sinápticos w em harmônicas, figuras 13 e 14, assim como nas verificações das outras bases, observa-se um padrão logarítmico e, sobre este padrão logarítmico, uma variação senoidal com decréscimo de acurácia para os casos de quantidade positiva de neurônios nas camadas ocultas. Observa-se também que acima de 120 pesos sinápticos a acurácia é regida majoritariamente pela primeira camada oculta (Onda0 e Onda1)

Aluno		RA/Matrícula		Professor	Tipo
Cristiano Lopes Moreira		119103-0		Dr Reinaldo Bianchi	Relatório de implementação
Data	Versão	Turma		Nome do arquivo	Página
12/12/2019	1	2º. Semestre de 2019		PEL_208_Relatório_06_Cristiano_Moreira.doc	31 (32)

5. Conclusão

Conclui-se que o MLP Perceptron é um algoritmo de redes neurais supervisionado de camada múltipla pertencente a arquitetura conectada de forma direta, não circular, feedforward, que utiliza a técnica de aprendizado pela propagação dos erros da saída aos elementos das camadas ocultas, back-propagation, que tem por objetivo classificar grupos sejam eles linearmente separáveis ou não.

O algoritmo comprova que a propagação dos erros para as camadas ocultas de uma rede neural artificial proposta por Rumelart, Hilton e Williams, possibilita o aprendizado nessas camadas e que as regras para a quantidade máxima de camadas ocultas e neurônios propostas por Hecht-Nielsen são coerentes com os experimentos realizados, porém, mesmo observando-se um comportamento logarítmico, com componentes de um “ruído” senoidal, na relação entre a quantidade de pesos sinápticos e a acurácia da RNA, ainda é um desafio acertar a quantidade ideal de camadas neurais, a quantidade ideal de neurônios e os ajustes finos de seus hiperparâmetros (taxa de aprendizado, número de épocas, função de ativação, quantidade de padrões...); sendo ainda os métodos de obtenção destes parâmetros a intuição e/ou uso de força bruta com diversas verificações, o que mostra a necessidade de pesquisas futuras sobre este tema.

6. Referências

- [1] MINSKY, Marvin Lee; PAPERT, Seymour A. **Perceptrons: an introduction to computational geometry**. Cambridge, Massachusetts: Mit Press, 1988.
- [2] RUSSELL, Stuart; NORVING, Peter. **Artificial Intelligence: A Modern Approach**. 3. ed. River, Nj, Usa: Pearson, 2010.
- [3] HASTIE, Trevor; TIBSHIRANI, Robert; FRIEDMAN, Jerome. **The Elements of Statistical Learning: Data Mining, Inference, and Prediction**. 2. ed. Stanford: Springer, 2008.
- [4] HAYKIN, Simon. **Redes Neurais: princípios e práticas**. 2. ed. Ontario Canada: Bookman, 2008.

Aluno		RA/Matrícula		Professor	Tipo
Cristiano Lopes Moreira		119103-0		Dr Reinaldo Bianchi	Relatório de implementação
Data	Versão	Turma		Nome do arquivo	Página
12/12/2019	1	2º. Semestre de 2019		PEL_208_Relatório_06_Cristiano_Moreira.doc	32 (32)