



UNIVERSIDADE DO MINHO

Trabalho de Grupo – 3º Exercício

Licenciatura em Engenharia Informática

Sistemas de Representação de Conhecimento e Raciocínio

3º Ano, 2º Semestre

Ano letivo 2014/2015

67646 Bruno Barbosa

53690 João Monteiro

54811 Hélder Machado

67707 Tiago Cunha

Departamento de Informática

4710-057

Braga, Portugal

Maio 2015

RESUMO

Neste relatório são apresentadas e descritas todos os procedimentos necessários à resolução do terceiro exercício do trabalho prático.

O objetivo principal para esta terceira fase consiste no uso de sistemas sub-simbólicos para representar o conhecimento e o desenvolvimento de mecanismos de raciocínio, mais concretamente, na utilização de Redes Neurais Artificiais para a resolução de problemas.

Para além das funcionalidades sugeridas, foram ainda introduzidas outras funcionalidades e características no sistema não só ao nível das capacidades de representação do conhecimento como também ao nível das faculdades de raciocínio.

CONTEÚDO

Resumo	2
Introdução	4
Preliminares.....	5
Descrição do Trabalho e Análise de Resultados	6
DataSet	6
TrainSet.....	7
TestSet	7
Qualidade do modelo	7
Rede Neuronal Artificial	8
Análise de resultados.....	9
Capacidade de resposta da rede	10
Sete níveis de fadiga	11
Existência ou ausência de fadiga	13
Melhor escala de fadiga	15
Conclusões e sugestões	16
Referências	17

INTRODUÇÃO

Foi-nos proposto no âmbito da unidade curricular de Sistemas de Representação de Conhecimento e Raciocínio utilizar um sistema sub-simbólico, em particular, uma rede neuronal artificial, para a resolução de um problema de identificação do nível de fadiga. Serão estudados as três capacidades de resposta que se seguem:

- Identificar os sete níveis de fadiga (que apresentaremos mais à frente).
- Identificar a existência ou ausência de fadiga.
- Encontrar a melhor escala de identificação de fadiga.

A escala de fadiga que se segue foi desenvolvida pelo doutor Layne Perelli nas forças armadas dos Estados Unidos da América. Esta escala tem sete níveis, dos quais o nível i representa menos fadiga do que o nível $i + 1$ e é bastante utilizada por profissionais numa área específica da biométrica: a fadiga mental. Assim sendo, a escala divide-se em:

1. Totalmente bem.
2. Responsivo, mas não no pico.
3. Ok, normal.
4. Em baixo de forma/do normal, a sentir-se em baixo.
5. Sentido moleza, perdendo foco.
6. Muito difícil concentrar, meio tonto.
7. Incapaz de funcionar, pronto a desligar.

Os dados que nos vão permitir treinar a nossa rede e a partir desta, inferir os principais sintomas resultantes da fadiga mental, são dados globais retirados de experiências previamente realizadas. Os resultados serão obtidos, como iremos demonstrar no relatório, através de redes neurais artificiais desenvolvidas no RStudio.

PRELIMINARES

Na área da ciência da computação as redes neuronais artificiais (RNA) correspondem uma representação simplificada do sistema nervoso central dos seres humanos. São um tipo de estrutura extraordinariamente interligada por neurónios ou nodos que possuem a capacidade de aprender. Existem duas semelhanças entre as redes neuronais artificiais e o cérebro humano que são:

- O conhecimento é adquirido a partir de um ambiente, através de um processo de aprendizagem ou de treino.
- O conhecimento é armazenado nas conexões, frequentemente designados por ligações ou sinapses, entre os nodos.

Desta forma, uma rede neuronal artificial possui um poder computacional alicerçado em dois aspetos essenciais:

- É composta por diversas unidades de processamento pequenas e simples que executam em paralelo.
- Têm a capacidade de aprendizagem e generalização, ou seja, responde de forma adequada a novas situações através experiências passadas.

Contudo, isto não significa que as redes neuronais artificiais sejam resposta para qualquer problema pois estão ainda muito longe de se equipararem à arquitetura do cérebro humano. Assim, em muitos casos, estas necessitam de estar incluídas com outros sistemas ou outros paradigmas de modo a darem respostas aos problemas em que estão inseridos.

Neste trabalho prático, serão usadas redes do tipo *Feedforward* Multicamada (RFMC) que se caracteriza pelo facto de existirem, uma ou várias, camadas intermédias que intervêm de forma vantajosa entre a entrada e saída da rede. De notar, que neste tipo de rede todas as conexões são unidireccionais no sentido dos nodos de entrada para os nodos de saída.

No nosso caso, temos oito nodos de entrada que correspondem às biométricas¹ utilizadas na deteção de fadiga, duas camadas intermédias com sessenta e quarenta nodos escondidos, respetivamente e uma único nodo de saída que corresponde ao nível de fadiga de mental.

¹ A biometria é a ciência que estuda as características físicas e/ou comportamentais dos seres vivos.

DESCRIÇÃO DO TRABALHO E ANÁLISE DE RESULTADOS

DataSet

O *dataset* que usamos na realização deste trabalho foi obtido através dados recolhidos relativos à interação humano-computador a partir dos dispositivos físicos de *input* tais como, o rato e o teclado. De seguida, deixamos um pequeno excerto do mesmo.

KDT Mean	MA Mean	MV Mean	TBC Mean	DDC Mean	DMS Mean	AED Mean	ADMSL Mean	Fatigue Level	Task
0,685	0,036	0,042	-0,086	0,099	0,000	-0,007	0,016	4	Work
0,022	0,038	0,042	-0,055	-0,016	0,002	0,014	0,058	4	Work
0,002	0,047	0,048	-0,095	0,524	0,008	-0,024	0,113	5	Work
0,000	0,052	0,058	-0,082	0,008	-0,011	0,023	-0,011	5	Work
0,008	0,077	0,081	-0,078	0,047	0,019	0,061	0,084	5	Work
0,002	0,031	0,036	-0,101	-0,009	-0,005	-0,021	0,000	5	Work
0,002	-0,029	-0,019	-0,099	-0,060	-0,007	0,046	-0,037	5	Work
0,074	0,018	0,032	-0,063	-0,045	0,000	-0,031	-0,001	3	Work
...

Figura 1. DataSet

Uma nota importante, que diz respeito à normalização dos dados na gama de $[-1; 1]$.

Como podemos observar, as primeiras oito colunas estão associadas aos nodos de entrada da rede neuronal artificial e correspondem às biométricas que foram consideradas para a determinação do nível de fadiga. Por exemplo,

- **KDTMean** – tempo médio compreendido entre o momento em que a tecla é pressionada para baixo e o momento que é largada.
- **MAMean** – aceleração média do rato durante um movimento, em determinado instante.
- **MVMean** – velocidade média do rato durante um movimento, em determinado instante.
- **TBCMean** – intervalo entre dois *clicks* consecutivos entre dois eventos de MOUSE_UP e MOUSE_DOWN.
- **DDCMean** – período de tempo entre dois eventos MOUSE_UP consecutivos.
- Entre outros.

TrainSet

A partir do *DataSet* foram retirados cerca de dois terços dos dados de modo a que pudéssemos construir a nossa rede neuronal e treiná-la através deste. A rede neuronal artificial é portanto, treinada, isto é, realiza um processo de aprendizagem, graças a este *set* de treino retirado do *DataSet*. Esta aprendizagem ocorre devido a uma fórmula que servirá como fator determinante na previsão do nível de fadiga. Essa fórmula corresponde à soma de todas as biométricas anteriormente referidas.

TestSet

Por sua vez, o *TestSet* é o nosso *set* de teste, ou seja, é o *set* com o qual vamos poder avaliar a capacidade de resposta da nossa rede e avaliar também se a aprendizagem foi, ou não, bem-sucedida. Trata-se de um terço restante do *DataSet* principal, contudo, são consideradas apenas as colunas referentes aos dados biométricos, pois o objetivo deste *set* é estimar nível de fadiga tendo em conta os dados de entrada efetuando, de seguida, uma comparação com os valores atuais. Desta forma, dependendo dos resultados vamos poder observar se a previsão deste *set* de teste vai de acordo com os níveis de fadiga já tabulados e assim considerar que a aprendizagem foi bem realizada.

Qualidade do modelo

Como acabamos de referir na secção anterior, a comparação dos resultados dos níveis de fadiga do *set* de treino com os valores já tabulados, determina não só, a qualidade de aprendizagem da rede assim como a qualidade do modelo que para determinados dados biométricos estima um nível de fadiga aceitável, semelhante aos já tabulados. No entanto, recorreremos ainda a outras funcionalidades que nos possibilitaram apurar a qualidade do modelo. Por exemplo, começamos com a impressão da rede neuronal após a realização do treino, que nos indica logo de imediato, a percentagem do erro e o número de passos necessários para chegar à conclusão. Logicamente, pretendemos que tanto um como o outro sejam o mais baixo possível.

```
> print(net)
Call: neuralnet(formula = formula, data = trainset, hidden = c(40, 20), threshold
= 0.01)

1 repetition was calculated.

Error Reached Threshold Steps
1 0.03202599803 0.009364744072 68799
```

Figura 2. Impressão da rede neuronal artificial

Por último, seguimos depois para o desvio padrão resultante das comparações entre os valores dos níveis de fadiga estimados e os níveis de fadiga já tabulados. Esta medida é obtida recorrendo à função *Root Mean Square Error*.

```
> #calcular o RMSE
> #RMSE dá o desvio padrão do erro de previsão do modelo. Um valor menor indica mel
hor desempenho do modelo.
> rmse(c(testset$FatigueLevel),c(results$previsao))
[1] 2.096390257
```

Figura 3. Desvio padrão

Neste caso, o valor alcançado foi na ordem das duas unidades. O significado desse mesmo valor é explicado na seguinte expressão.

$$valor_{estimado} \in [valor_{tabulado} - desvio; valor_{tabulado} + desvio]$$

Rede Neuronal Artificial

Como já foi referido anteriormente, a rede criada para este trabalho prático é constituída por quatro camadas. Na primeira camada, tem oito nodos de entrada, que condizem com as biométricas. A segunda e terceira camadas são também designadas pelas camadas intermédias e têm quarenta nodos e vinte nodos, respetivamente. Na quarta e última camada, temos o tão desejado nível de fadiga. Este tipo de rede (*feedforward*) tem a particularidade de que as conexões entre as camadas é unidirecional, no sentido dos nodos de entrada para os nodos de saída.

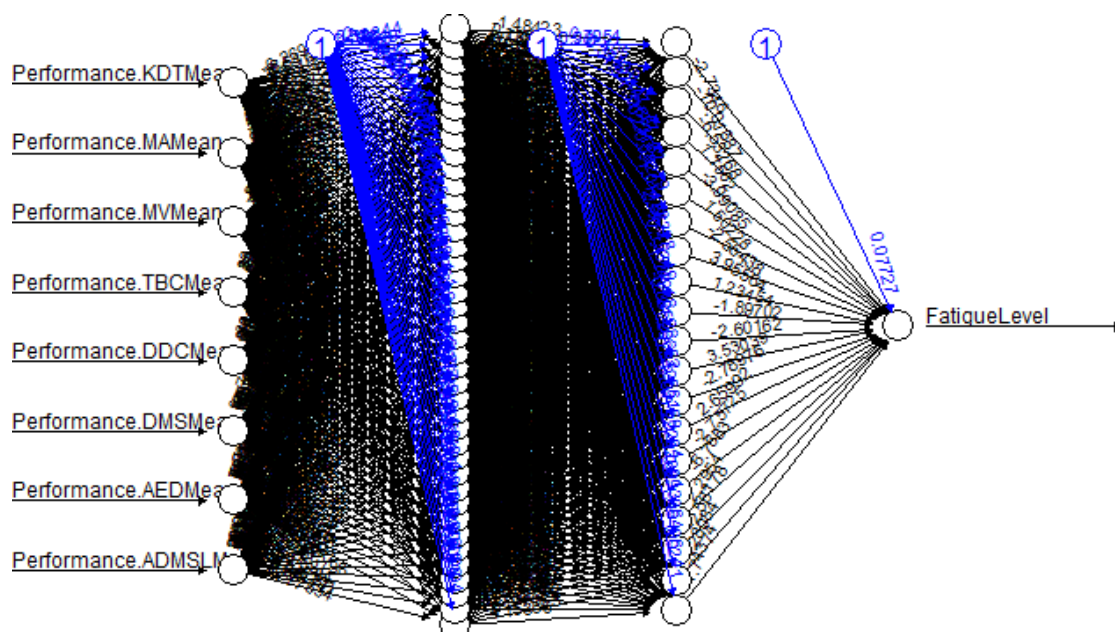


Figura 4. Rede neuronal artificial

Eis o comando usado em R para gerar a rede em questão. Repare que nesta rede o valor do campo *threshold* é de uma centésima. Este valor funciona como critério de paragem na aprendizagem, por isso, quanto menor for, mais exata é a rede resultante.

```
#Treinar a rede neural
#vai ter 40 e 20 camadas ocultas
#Threshold é um valor numérico que especifica o limite para o erro parcial
net <- neuralnet(formula,trainset, hidden=c(40,20), threshold=0.01)
```

Figura 5. Criação da rede neuronal artificial

Considere que esta foi a nossa rede base. Na próxima secção apresentaremos diversos resultados obtidos variando algumas das características da rede, assim como o tamanho dos *sets* de treino e de teste, de modo a que a rede seja mais eficiente e mais precisa ao estimar o nível de fadiga.

Análise de resultados

Nesta secção apresentamos as várias experiências que fizemos. Estas experiências permitiram-nos, passo a passo, melhorar a nossa rede neuronal em termos de eficiência e precisão. A tabela que se segue demonstra, para cada rede, o número de nodos escondidos nas camadas intermédias, o valor do erro, o número de passos necessários para a aprendizagem e o desvio padrão.

Nestes primeiros resultados o *set* de treino tem 600 elementos, enquanto o *set* de teste tem os restantes elementos do *set* principal, ou seja, 244 elementos. Os resultados foram os seguintes.

Nodos Intermédios	Erro	Passos	Desvio
(60,40)	0.0043	6541	1.79
(55,40)	0.0024	13468	2.26
(60,30)	0.00662	18014	1.81
(40,30)	0.00299	16573	2.16
(60,50)	0.00181	8093	1.82
(70,50)	0.00193	11437	1.79
(70,60)	0.0025	7157	1.76
(80,60)	0.0022	8593	1.85
(40,20)	0.0140	44082	2.41
(80,70)	0.0011	5113	1.73

Tabela 1. Resultados segundo os nodos intermédios

Uma vez analisados os valores da tabela 1, a principal preocupação foi escolher o número de nodos intermédios onde o erro fosse na ordem de uma milésima e o desvio padrão o mais inferior possível. Sendo assim, optamos por oitenta nodos na primeira camada intermédia e setenta nodos na segunda camada. Para estes valores voltámos a efetuar outros testes, agora desta vez, variando os tamanhos dos *sets* de treino e de teste e comparando os tempos de aprendizagem.

TrainSet	TestSet	User	System	Elapsed
600	244	148.67	0.01	148.75
500	344	115.56	0.01	115.75
400	444	53.66	0.00	53.72

Tabela 2. Tempos de aprendizagem

Como podemos observar, à medida que seleccionamos um *TrainSet* maior, o tempo de aprendizagem sobe substancialmente. Todavia, a aprendizagem é mais rica e a rede consegue ser mais nítida na determinação dos níveis de fadiga. Por questões de consistência elegemos o *TrainSet* de 600 elementos para que a aprendizagem fosse mais robusta.

Capacidade de resposta da rede

Relembrando que, com este trabalho prático, deseja-se que sejam,

- Identificados os sete níveis de fadiga tal como descrito pelos dados.
- Identificada, simplesmente, a existência ou ausência de fadiga.
- Encontrada a melhor escala de identificação de fadiga.

No primeiro ponto, consideramos que depois da aprendizagem o sistema tem de ser capaz de estimar o nível de fadiga de acordo com os dados de teste. De notar que a previsão não resulta num valor inteiro pelo que, desta forma, esse valor foi arredondado sem casas decimais para que respeitasse ao máximo a escala previamente exposta.

No segundo ponto, consideramos que a existência de fadiga surge a partir do nível quatro da escala visto que no nível anterior ainda é caracterizado como sendo “*Ok, normal*” enquanto o quarto nível já é “*Em baixo de forma/do normal, a sentir-se em baixo*”.

O terceiro e último ponto foi o que gerou mais dúvidas sobre o que era para se fazer. Mas depois de esclarecido, foi bem mais fácil do que estávamos a pensar. A nossa sugestão para a escala fadiga é inserindo dez níveis entre os níveis já existentes. Por isso, em vez de termos uma escala em unidades, passamos a ter uma escala com uma casa decimal e desta forma, conseguimos ser mais precisos na detenção da transição do estado de ausência de fadiga para a presença de fadiga.

Sete níveis de fadiga

	atual	previsao						
601	3	2	637	3	1	674	2	1
602	1	3	638	3	4	675	2	5
603	1	0	639	3	3	676	2	2
604	2	2	640	3	1	677	2	1
605	2	1	641	3	3	678	2	4
606	1	5	642	3	3	679	2	4
607	1	1	643	3	2	680	3	3
608	3	2	644	3	1	681	3	4
609	3	2	645	3	4	682	3	2
610	4	2	646	3	3	683	3	2
611	4	4	647	3	2	684	3	1
612	3	2	648	3	4	685	3	3
613	3	5	649	1	2	686	3	4
614	3	3	650	1	2	687	3	4
615	3	6	651	1	3	688	3	3
616	3	4	652	1	2	689	3	2
617	3	4	653	1	3	690	4	3
618	3	6	654	1	3	691	4	3
619	2	2	655	3	2	692	4	3
620	2	3	656	3	5	693	4	3
621	2	1	657	3	0	694	5	3
622	2	4	658	3	1	695	5	3
623	2	5	659	4	1	696	4	4
624	2	3	660	4	2	697	4	5
625	4	4	661	4	4	698	4	0
626	4	2	662	4	2	699	4	2
627	4	2	663	4	3	700	4	2
628	4	3	664	4	3	701	4	4
629	4	3	665	4	6	702	4	4
630	4	6	666	4	3	703	4	3
631	4	2	667	4	0	704	3	4
632	4	3	668	4	3	705	3	3
633	3	3	669	4	3	706	3	1
634	3	3	670	4	2	707	3	3
635	3	2	671	2	7	708	3	2
636	3	1	672	2	3	709	3	1
			673	2	5	710	3	2

Figura 6. Previsão do nível de fadiga

Pela figura 6, podemos observar as previsões obtidas. São visíveis alguns casos fora de escala como por exemplo, o do elemento 603 do *set* principal, onde a previsão foi de zero (fora da escala!) e o valor atual é de um. Este efeito acontece, como já foi referido, devido ao desvio padrão. De qualquer maneira, na maioria dos casos os resultados estão dentro da escala, tendo ainda algumas previsões corretas.

Nesta página vamos apresentar os resultados obtidos, agora considerando apenas cinco das oitos biométricas. Inesperadamente, o tempo de aprendizagem foi significativamente superior atingindo os 766 segundos, que correspondem aproximadamente a 13 minutos. Em termos de medidas de qualidade da rede os resultados foram razoáveis, com um erro de 0.0143 em 29134 passos. O desvio padrão foi de 2.1434.

	atual previsao		636	3	4	672	2	4
601	3	3	637	3	0	673	2	3
602	1	1	638	3	2	674	2	3
603	1	5	639	3	5	675	2	3
604	2	1	640	3	2	676	2	3
605	2	1	641	3	2	677	2	2
606	1	1	642	3	5	678	2	3
607	1	4	643	3	5	679	2	3
608	3	7	644	3	0	680	3	3
609	3	6	645	3	4	681	3	3
610	4	6	646	3	3	682	3	2
611	4	3	647	3	2	683	3	2
612	3	2	648	3	3	684	3	1
613	3	-3	649	1	3	685	3	2
614	3	3	650	1	3	686	3	3
615	3	1	651	1	1	687	3	2
616	3	1	652	1	1	688	3	2
617	3	0	653	1	9	689	3	3
618	3	3	654	1	0	690	4	4
619	2	3	655	3	5	691	4	1
620	2	3	656	3	3	692	4	3
621	2	2	657	3	2	693	4	3
622	2	1	658	3	1	694	5	3
623	2	3	659	4	2	695	5	4
624	2	5	660	4	5	696	4	4
625	4	5	661	4	4	697	4	4
626	4	3	662	4	5	698	4	2
627	4	0	663	4	3	699	4	3
628	4	3	664	4	5	700	4	2
629	4	4	665	4	5	701	4	4
630	4	4	666	4	3	702	4	4
631	4	7	667	4	4	703	4	6
632	4	5	668	4	4	704	3	4
633	3	3	669	4	2	705	3	2
634	3	5	670	4	5	706	3	1
635	3	5	671	2	4	707	3	3

Figura 7. Resultados com escala normal e 5 atributos

Existência ou ausência de fadiga

Nesta situação começamos por criar uma cópia do *dataset* principal. Posto isto, o próximo passo consistiu em alterar a escala de fadiga de um a sete, para zero ou um, segundo a expressão que se segue.

$$\begin{cases} 0 & \text{se o nível de fadiga} \leq 3 \\ 1 & \text{se o nível de fadiga} > 3 \end{cases}$$

```
#header -> é uma variável lógica que indica se o ficheiro tem ou não os nomes das variáveis na primeira linha.  
#sep -> é o carácter que separa os valores em cada linha  
#dec -> é o carácter utilizado como separador decimal.  
data <- read.csv("C:\\Users\\Tiago Cunha\\Desktop\\TP3\\exercício3.csv", header=TRUE, sep=";", dec=".")  
  
newdata <- data  
  
#Transformar fadiga em 0 ou 1  
newdata$FatigueLevel[newdata$FatigueLevel <= 3] <- 0  
newdata$FatigueLevel[newdata$FatigueLevel > 3] <- 1
```

Figura 8. Transformação da escala de fadiga

Eis os resultados que obtivemos. Se compararmos com os resultados obtidos com a escala normal (figura 6) vemos que os níveis de fadiga anteriores condizem com a nova escala. Porém, foram indispensáveis reajustes no número de camadas e no número de nodos escondidos porque os valores de erro do modelo, o número de passos de aprendizagem e o desvio padrão tomaram valores relativamente superiores. Por exemplo, o desvio padrão, pela primeira vez que realizou a aprendizagem, rondou as sete décimas, o que numa escala como esta é muito elevado. Eis o erro relativo nas duas escalas.

$$(escala de 1 a 7) \frac{1.73}{7} = 0.25 \text{ (2 c. d.)} < \frac{0.73}{1} = 0.73 \text{ (escala de 0 a 1)}$$

Após de várias experiências rematamos por considerar para esta nova escala, uma rede neuronal com cinco camadas. A já conhecida camada de entrada com os oito nodos de entrada correspondentes às biométricas, três camadas intermédias com, respetivamente, 50,35 e 20 nodos escondidos e a última camada, a camada de saída, referente ao nível de fadiga.

Na próxima página demonstramos o melhor resultado obtido depois após inúmeros testes. Verificou-se evidentemente, a grande percentagem de erro da estimativa face ao valor tabulado. Prova disso são as previsões fora da escala, muitas vezes, negativas ou até mesmo superiores a um.

			637	0	0			
	atual	previsao	638	0	0	674	0	0
601	0	0	639	0	0	675	0	0
602	0	-1	640	0	0	676	0	1
603	0	0	641	0	1	677	0	0
604	0	0	642	0	0	678	0	0
605	0	0	643	0	0	679	0	1
606	0	1	644	0	1	680	0	1
607	0	0	645	0	0	681	0	0
608	0	-1	646	0	0	682	0	0
609	0	0	647	0	0	683	0	0
610	1	0	648	0	0	684	0	1
611	1	0	649	0	0	685	0	0
612	0	1	650	0	0	686	0	1
613	0	0	651	0	0	687	0	0
614	0	0	652	0	0	688	0	1
615	0	0	653	0	1	689	0	0
616	0	0	654	0	0	690	1	1
617	0	0	655	0	0	691	1	0
618	0	0	656	0	0	692	1	0
619	0	0	657	0	0	693	1	0
620	0	0	658	0	0	694	1	0
621	0	0	659	1	-1	695	1	0
622	0	1	660	1	0	696	1	1
623	0	0	661	1	0	697	1	1
624	0	1	662	1	0	698	1	0
625	1	0	663	1	0	699	1	0
626	1	0	664	1	0	700	1	0
627	1	0	665	1	0	701	1	0
628	1	0	666	1	-1	702	1	0
629	1	0	667	1	0	703	1	0
630	1	0	668	1	0	704	0	0
631	1	1	669	1	0	705	0	0
632	1	0	670	1	0	706	0	1
633	0	0	671	0	1	707	0	0
634	0	0	672	0	1	708	0	0
635	0	0	673	0	0	709	0	0
636	0	1				710	0	0

Figura 9. Previsões com a escala binária

De seguida, temos a figura 9, como prova do desvio padrão da estimativa.

```
> #calcular o RMSE
> #RMSE dá o desvio padrão do erro de previsão do modelo.
> rmse(c(testset$FatigueLevel),c(results$previsao))
[1] 0.6559938024
> |
```

Figura 10. Desvio padrão da estimativa

Melhor escala de fadiga

Perante uma situação de identificar a melhor escala de fadiga, o nosso grupo considerou que, para que fossemos mais exatos a indicar quando ocorre a transição do estado de ausência de fadiga para o estado de existência de fadiga, deveriam existir dez níveis intermédios entre cada nível da escala principal (sete níveis). Deste modo, de acordo com o que foi dito na parte da identificação da existência/ausência de fadiga acontecia entre os níveis três e quatro. Portanto, com a nossa sugestão de escala o ponto de decisão sucede no nível 3.5. Em termos de medidas de precisão o desvio foi de 1.77 e o erro da rede 0.00111 com 5113 passos em 385.13 segundos.

	atual	previsao						
601	3	1.8	637	3	1.3	674	2	1.5
602	1	2.9	638	3	3.8	675	2	5.4
603	1	-0.3	639	3	3.0	676	2	2.4
604	2	1.9	640	3	1.0	677	2	1.1
605	2	0.6	641	3	2.8	678	2	3.5
606	1	4.6	642	3	2.8	679	2	4.5
607	1	0.6	643	3	1.8	680	3	2.9
608	3	1.8	644	3	1.2	681	3	4.3
609	3	2.1	645	3	4.2	682	3	1.7
610	4	2.3	646	3	3.0	683	3	2.3
611	4	3.9	647	3	2.0	684	3	1.1
612	3	1.8	648	3	3.9	685	3	3.3
613	3	4.8	649	1	1.7	686	3	4.3
614	3	2.5	650	1	2.1	687	3	4.4
615	3	5.5	651	1	2.8	688	3	3.2
616	3	3.6	652	1	1.9	689	3	2.0
617	3	3.7	653	1	2.8	690	4	2.7
618	3	6.3	654	1	2.7	691	4	3.3
619	2	2.4	655	3	1.8	692	4	3.3
620	2	3.3	656	3	5.3	693	4	2.6
621	2	0.7	657	3	0.1	694	5	2.7
622	2	4.4	658	3	0.8	695	5	3.3
623	2	4.9	659	4	1.2	696	4	3.9
624	2	2.7	660	4	2.3	697	4	4.8
625	4	4.2	661	4	4.2	698	4	-0.4
626	4	2.5	662	4	1.8	699	4	1.8
627	4	2.1	663	4	3.3	700	4	2.0
628	4	2.5	664	4	2.8	701	4	3.8
629	4	2.9	665	4	5.8	702	4	4.5
630	4	6.0	666	4	3.0	703	4	3.3
631	4	2.2	667	4	0.5	704	3	4.4
632	4	3.4	668	4	3.1	705	3	2.6
633	3	2.7	669	4	3.3	706	3	1.5
634	3	2.8	670	4	2.5	707	3	3.1
635	3	2.3	671	2	6.7	708	3	2.5
636	3	1.1	672	2	3.1	709	3	1.2
			673	2	4.8	710	3	2.4

Figura 11. Melhor escala para identificar fadiga

CONCLUSÕES E SUGESTÕES

No final deste terceiro exercício do trabalho de grupo proposto, consideramos que foram atingidos os objetivos essenciais à resolução do mesmo.

Terminado o trabalho, podemos concluir que a escala principal é preferível em relação às outras escalas sugeridas. E os fatores decisivos nesta nossa decisão foram os resultados da função *Root Mean Square Error* e erro resultante da criação e aprendizagem da rede. Embora também pudéssemos considerar o tempo que o sistema demorou a treinar a rede, não quisemos ver isso como uma condição, preferindo desta forma, dar mais relevância à capacidade de previsão.

Como tal, o rácio de desvio padrão e a escala escolhida é relativamente baixo, pelo que é possível prever determinado nível de fadiga tendo em conta os dados biométricos recebidos como *input*. Posto isto, concluímos também que o tamanho do *set* de treino é fundamental para a robustez da rede isso torna a aprendizagem mais rica e melhora a capacidade de esta estimar um valor mais próximo do esperado.

Não querendo deixar de referir, que foram realizados vários ensaios onde o número de nodos escondidos nas camadas intermédias eram próximos de 16 e 8 nodos, respetivamente. Todavia, os valores referentes à qualidade da rede estavam muito longe de serem desejados, nesse sentido, preferimos colocar um maior número de nodos escondidos nas camadas intermédias da rede.

Em suma, achamos que o trabalho foi bem-sucedido e que dos pontos propostos, foram todos razoavelmente bem atingidos.

REFERÊNCIAS

- [1] BRATKO, Ivan
“Prolog Programming for Artificial Intelligence, 4th Edition”
Addison-Wesley, University of Ljubljana, 24 de Agosto de 2011

- [2] COELHO, Hélder
“A Inteligência Artificial em 25 Lições”
Fundação Calouste Gulbenkian, 1995

- [3] CORTEZ, Paulo, NEVES, José
“Redes Neurais Artificiais”
Universidade do Minho, Escola de Engenharia, Departamento de Informática, Braga, 2000

- [4] ANALIDE, César, NOVAIS, Paulo, NEVES, José
“Redes Neurais Artificiais”
Universidade do Minho, Escola de Engenharia, Departamento de Informática, ISLab

- [5] PIMENTA, André, ANALIDE, César, NOVAIS, Paulo
“Data analysis using R”
Universidade do Minho, Escola de Engenharia, Departamento de Informática, ISLab