

CPE 721 – 2012 - RNs Feedforward
2ª Série de Exercícios - Pré-processamento

1 – Considere as variáveis x_1, x_2 e y , todas com média nula e desvio padrão σ_1, σ_2 e σ_y , reespectivamente. r_{12} é a correlação (coeficiente de correlação de Pearson) entre x_1 e x_2 , r_{1y} é a correlação entre x_1 e y , e r_{2y} é a correlação entre x_2 e y . A variável $z = x_2 - ax_1$ é a variável x_2 “descorrelacionada” da variável x_1 , e r_{zy} é a correlação entre z e y . O parâmetro a é obtido pela minimização de $E(x_2 - ax_1)^2$. Prove que:

$$a = \frac{\sigma_2}{\sigma_1} r_{12} \quad \text{e} \quad r_{zy} = \frac{r_{2y} - r_{1y} r_{12}}{\sqrt{1 - r_{12}^2}}$$

2 – Apresente a normalização a ser usada para os pares entrada-saída ($\underline{X}$, Y) da Tabela 1 abaixo.

2.1 - Use correlação para verificar se alguma variável de entrada pode ser eliminada.

2.2 - Verifique a presença de possíveis intrusos. Esta verificação é mais fácil após a normalização (mas também é possível com os dados brutos)

Para quem teve dificuldade em conseguir os programas de análise estatística seguem abaixo algumas tabelas (valores aproximados) para auxiliar a resolução do exercício 2. Considere que as distribuições das variáveis podem ser aproximadas por gaussianas.

Variável	X1	X2	X3	X4	Y
Média	9,73	-0,95	2,11	5,01	3,44
D.P.	1,77	1,14	0,64	1,34	1,84

Correlação	X1	X2	X3	X4	Y
X1	1,00	0,01	0,07	-0,59	-0,64
X2		1,00	-0,20	-0,04	0,17
X3			1,00	0,10	-0,59
X4				1,00	0,62

Tabela 1

Par	X1	X2	X3	X4	Y
1	7,73	-1,57	1,38	4,56	4,62
2	9,19	-1,80	2,84	4,58	2,20
3	9,32	-1,25	2,63	6,63	3,49
4	7,40	-0,64	1,47	6,19	5,14
5	12,49	-0,07	1,09	2,93	3,24
6	8,18	-2,01	2,47	5,87	4,21
7	9,75	-2,50	1,48	4,98	4,67
8	9,11	-2,91	2,93	5,99	2,92
9	9,72	-1,77	1,33	3,37	3,59
10	7,06	-0,53	2,33	5,14	4,23
11	4,59	-1,29	2,74	5,32	3,15
12	9,13	-0,30	2,82	6,83	3,67
13	10,62	-0,28	1,05	5,83	8,97
14	13,10	-1,28	2,95	3,24	-0,82
15	13,40	-1,2	1,27	2,12	0,95
16	6,85	-1,67	2,15	7,06	4,67
17	13,11	0,22	2,77	4,05	0,29

18	9,59	0,38	2,69	4,48	2,15
19	11,53	-0,05	2,58	4,01	1,64
20	8,51	0,79	1,97	4,91	4,83
21	10,22	-2,98	2,64	4,98	4,43
22	9,43	-1,27	2,06	4,41	3,79
23	10,93	-1,30	1,95	4,75	2,50
24	9,72	-2,02	2,31	3,93	2,22
25	12,39	-1,32	1,77	5,33	2,41
26	9,98	-2,72	1,97	6,23	4,45
27	7,37	0,50	1,64	4,90	4,99
28	13,07	-1,30	1,31	1,56	0,76
29	10,57	0,14	2,79	5,28	1,58
30	8,92	-2,41	1,79	6,38	4,28
31	12,26	-1,69	2,48	3,18	0,85
32	7,49	-0,24	2,62	7,13	4,72
33	9,07	-0,10	2,55	7,40	5,03
34	6,89	0,03	1,20	4,89	5,71
35	8,72	0,26	1,39	5,34	4,70
36	10,38	0,94	1,71	3,19	3,54

3 – Projete uma rede neural para realizar o mapeamento da Tabela 1 acima. Aloque 20 pares para treinamento, 8 para teste e 8 para validação. Os possíveis intrusos, se existirem, devem ser alocados no conjunto de teste, e os demais pares sorteados pelos três conjuntos. Projete a rede usando apenas as variáveis que persistiram após a eliminação usando correlação linear. Não é necessário fazer a validação cruzada, mas faça o acompanhamento durante o treinamento e a crítica pós treinamento. Sugestão: utilize uma rede com duas camadas de neurônios, com 4 neurônios na camada intermediária.

3.1 - Repita inserindo uma a uma as variáveis eliminadas pelo critério de correlação linear. Compare os resultados obtidos.

Preserve os resultados, serão necessários para exercícios posteriores.