



EPC 3

- 1- Treine e valide a melhor topologia usando 10-folds cross-validation. Execute três vezes cada fold usando o algoritmo de aprendizagem backpropagation padrão, inicializando as matrizes de pesos com valores aleatórios entre 0 e 1. Utilize a função de ativação logística (sigmoid) para todos os neurônios, taxa de aprendizado $\eta = 0.1$, precisão $\varepsilon = 10^{-6}$ e $\beta = 0.5$. Para cada topologia testada, registre em uma tabela a média e o desvio padrão das seguintes medidas: EQM, Número de épocas, Tempo (segundos) e Acurácia (porcentagem de acerto do conjunto de validação).

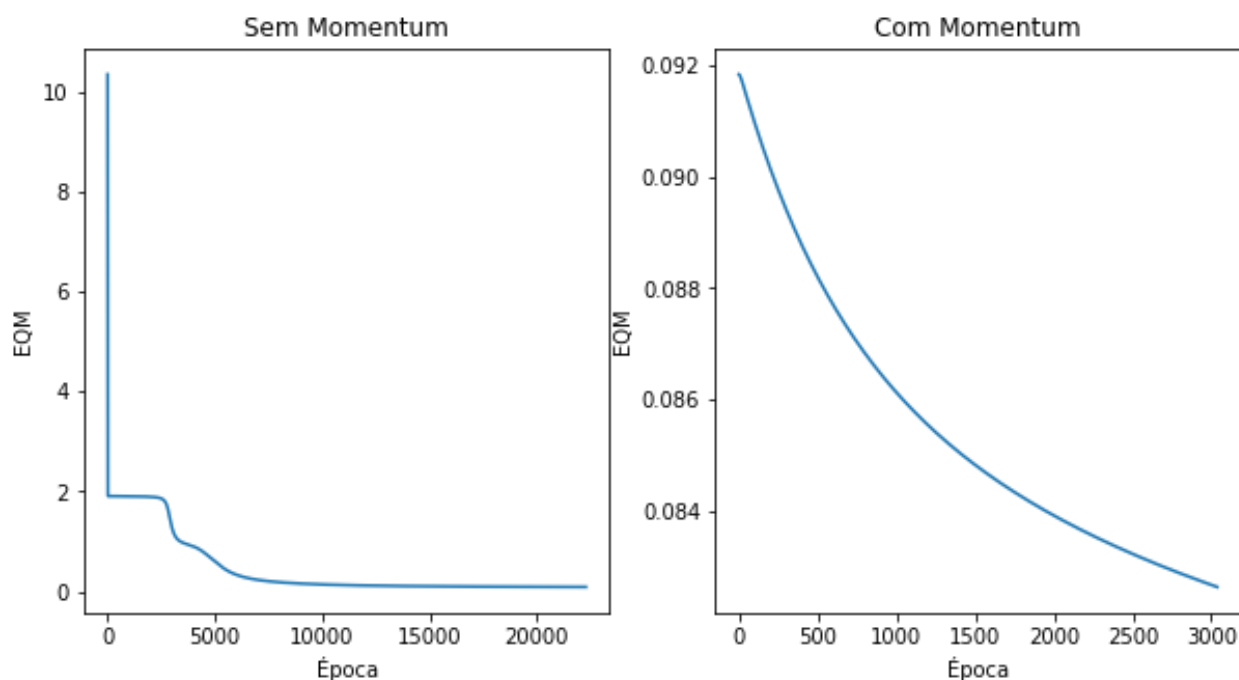
Nº de Neurônios	EQM		Épocas		Tempo		Acurácia	
	Média	DP	Média	DP	Média	DP	Média	DP
(04,04)	1.48171	0.77257	7788.5	13924.98	2.52	4.52	0.37142	0.31965
(05,05)	0.80356	0.90224	19088.93	15921.48	6.07	5.06	0.65	0.38036
(07,07)	0.1309	0.33133	29128.16	7817.49	9.75	2.60	0.92857	0.15319
(09,09)	0.19900	0.45733	26001.53	9031.75	9.03	3.11	0.92857	0.18442
(04,10)	0.143	0.32852	29124	7368.68	9.66	2.42	0.93809	0.14253
(05,10)	0.20786	0.45481	25438.36	7678.33	8.47	2.56	0.94047	0.19525
(06,10)	0.06918	0.03099	29314.36	4905.33	9.97	1.66	0.9619	0.06317
(07,10)	0.07869	0.0226	27100.16	3939.45	9.38	1.37	0.96428	0.04791
(08,10)	0.13932	0.33007	26256.23	6480.91	9.13	2.26	0.94761	0.12768
(09,10)	0.13787	0.33032	25836.2	6434.50	9.09	2.26	0.93095	0.14461
(10,10)	0.25363	0.55193	26940.46	10452.27	9.54	3.69	0.89523	0.23742
(11,10)	0.26809	0.54712	24632.56	9505.07	8.86	3.42	0.91428	0.22692
(10,09)	0.50	0.77762	23508.43	14294.82	8.29	5.03	0.80714	0.30647
(10,11)	0.19710	0.45844	27020.53	9600.18	9.74	3.45	0.91666	0.18724
(10,12)	0.20297	0.4565	24840.16	8130.45	8.98	2.93	0.91904	0.19231
(20,10)	1.9073	0.00176	45.13	0.76303	0.01	0.0013	0.2	0.04285

(i, j) representa i neurônios na 1ª camada escondida e j na 2ª

- 2- Após encontrar a melhor topologia execute três vezes cada fold através do algoritmo de aprendizagem backpropagation com momentum, utilizando as mesmas matrizes de pesos iniciais que foram usadas no item (1). Utilize a função de ativação logística (sigmoid) para todos os neurônios, taxa de aprendizado $\eta=0.1$, fator de momentum $\alpha=0.9$, precisão $\varepsilon = 10^{-6}$ e $\beta = 0.5$. Registre em uma tabela a média e o desvio padrão das seguintes medidas: EQM, Número de épocas, Tempo (segundos) e Acurácia (porcentagem de acerto do conjunto de validação).

Nº de Neurônios	EQM		Épocas		Tempo		Acurácia	
	Média	DP	Média	DP	Média	DP	Média	DP
(07,10)	0.15798	0.57105	5626.06	3214.78	1.89	1.08	0.94285	0.07

- 3- Para a melhor topologia encontrada, de cada PMC (treinada com backpropagation padrão e com momentum), trace os respectivos gráficos dos valores de erro quadrático médio (EQM) em função de cada época de treinamento. Meça também o tempo de processamento. Imprima os gráficos lado a lado, para facilitar a comparação visual entre as redes.



Tempo sem Momentum: 7.98082423210144

Tempo com Momentum: 1.0419366359710693

- 4- Dado que o problema se configura como um típico processo de classificação de padrões, implemente uma rotina que faz o pós-processamento das saídas fornecidas pela rede (números reais) para números inteiros. Utilize o critério do arredondamento simétrico.
- 5- Teste as redes usando os conjuntos de teste de cada problema. Registre em uma tabela a média e o desvio padrão das taxas de acerto (%) em cada problema.

	Média	Desvio Padrão
Sem Momentum	93,57%	0,05933
Com Momentum	91,42%	0.05345

- 6- Explique o que são situações de underfitting e overfitting, descrevendo-se também os meios para as suas detecções. Quais são as possíveis soluções para cada situação?

R: Underfitting: sub ajustado, ou seja, a rede apresenta muitos erros tanto no treinamento como no teste. Com isso a rede não conseguiu aprender o suficiente.

Overfitting: sobre ajustado, ou seja, a rede conseguiu aprender muito com o treinamento, porém ela aprendeu em excesso e com isso apenas decorou. Ao apresentar novos dados para o teste a rede vai errar muito pois ela aprendeu apenas a identificar os dados do treinamento.

Ambos os casos podem ser identificados utilizando uma visualização gráfica ou através da validação cruzada ao medir a acurácia da rede para a resolução dos problemas.

O underfitting pode ser resolvido aumentando o número de camadas ou neurônios e utilizar dados que possuem uma boa relação entre si e o que queremos como resposta. Já o overfitting pode ser resolvido simplificando o modelo, utilizar mais dados para o treinamento e utilizar um pré-processamento para tratar os dados e eliminar ruídos.