



## Respostas

### 1) Questão:

Utilização do dataset: fashion\_mnist

GitHub: [https://github.com/paulovpcotta/trabalho2\\_unb\\_ml](https://github.com/paulovpcotta/trabalho2_unb_ml)

Jupyter: Trabalho\_ML\_Paulo\_Cotta - Q1

Foi implementado um Multilayer Perceptron (MLP) usando o framework keras para solucionar um problema de classificação multiclases, usando a base de dados MNIST.

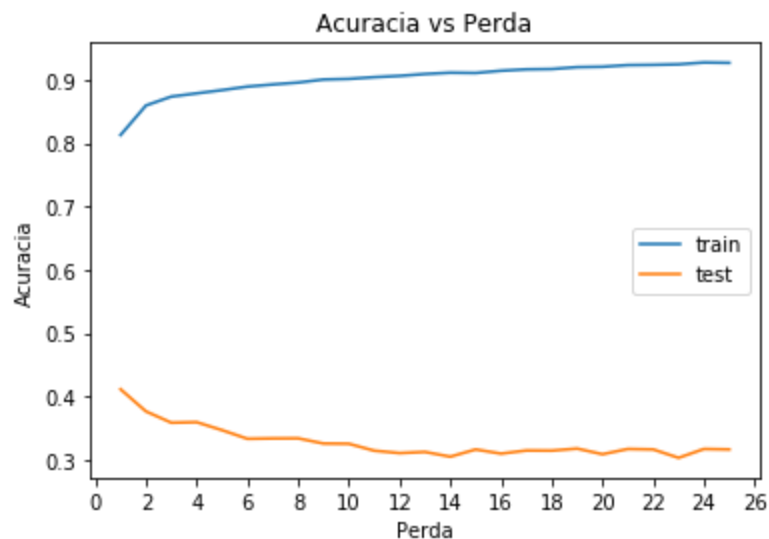
Foi usado o dataset do Fashion Mnist, para representar o problema proposto.

Segue o sumário da rede neural.

Model: "sequential\_1"

Layer (type)	Output Shape	Param #
dense_1 (Dense)	(None, 512)	401920
activation_1 (Activation)	(None, 512)	0
dropout_1 (Dropout)	(None, 512)	0
dense_2 (Dense)	(None, 512)	262656
activation_2 (Activation)	(None, 512)	0
dropout_2 (Dropout)	(None, 512)	0
dense_3 (Dense)	(None, 10)	5130
activation_3 (Activation)	(None, 10)	0
Total params: 669,706		
Trainable params: 669,706		
Non-trainable params: 0		

Segue performance do modelo:



Melhor acurácia: 93%

Matriz de confusão

	Atual	T-shirt/top	Trouser	Pullover	Dress	Coat	Sandal	Shirt	Sneaker	Bag	Ankle boot
Preditivo											
T-shirt/top	846	0	13	19	2	1	118	0	1	0	0
Trouser	4	973	0	16	4	0	3	0	0	0	0
Pullover	22	1	836	6	72	0	63	0	0	0	0
Dress	27	2	14	889	34	0	32	0	2	0	0
Coat	3	0	130	25	788	0	54	0	0	0	0
Sandal	0	0	0	0	0	974	0	18	0	0	8
Shirt	123	0	84	21	64	0	707	0	1	0	0
Sneaker	0	0	0	0	0	5	0	977	0	0	18
Bag	7	0	2	5	3	1	13	3	966	0	0
Ankle boot	1	0	0	0	0	6	0	39	0	0	954

## 2) Questão

Utilização do dataset: fashion\_mnist

GitHub: [https://github.com/paulovpcotta/trabalho2\\_unb\\_ml](https://github.com/paulovpcotta/trabalho2_unb_ml)

Jupyter: Trabalho\_ML\_Paulo\_Cotta - Q2

Foi usado Mnist Fashion para tratar o mesmo problema.

Usando as labels:

labels = {

```
0:"T-shirt/top",
1:"Trouser",
2:"Pullover",
3:"Dress",
4:"Coat",
5:"Sandal",
6:"Shirt",
7:"Sneaker",
8:"Bag",
9:"Ankle boot"
}
```

Executando o predict para classificação, não deu uma boa acurácia, sendo que o modelo não atingiu 52% de assertividade.

### 3) Questão

Utilização do dataset: Pattern Random

GitHub: [https://github.com/paulovpcotta/trabalho2\\_unb\\_ml](https://github.com/paulovpcotta/trabalho2_unb_ml)

Jupyter: Trabalho\_ML\_Paulo\_Cotta - Q3

O algoritmo foi implementado na classe do Jupyter (RBFNetwork).

Total de acurácia: 50%

Last MSE: 0.21317541752461472

### 4) Questão

Utilização do dataset: Pima Indians Diabetes Database

GitHub: [https://github.com/paulovpcotta/trabalho2\\_unb\\_ml](https://github.com/paulovpcotta/trabalho2_unb_ml)

Jupyter: Trabalho\_ML\_Paulo\_Cotta - Q4

Este conjunto de dados é originalmente do National Institute of Diabetes and Digestive and Kidney

Diseases. O objetivo do conjunto de dados é diagnosticar se um paciente tem ou não diabetes, com

base em determinadas medidas de diagnóstico incluídas no conjunto de dados. Várias restrições

foram colocadas na seleção dessas instâncias de um banco de dados maior. Em particular, todos os

pacientes aqui são do sexo feminino, pelo menos, 21 anos de idade de origem indiana Pima. São

768 amostras.

Atributo de saída: positivo ou negativo para diabetes.

500 amostras negativas e 268 positivas para diabetes.

Informações sobre atributos:

1. número de vezes grávida;

2. concentração plasmática de glicose 2 horas em teste oral de tolerância à glicose;
3. pressão arterial diastólica (mmHg);
4. espessura da dobra da pele do tríceps (mm);
5. insulina sérica de 2 horas ( $\mu\text{U} / \text{ml}$ );
6. índice de massa corporal (peso em kg / (altura em m) <sup>2</sup>);
7. função de pedigree de diabetes;
8. idade (anos);
9. variável da classe (0 – negativo ou 1 – positivo);

Kernel Linear	
C	Erro de classificação (média dos folds)
0,25	23,0469 %
0,5	22,6563 %
1	22,6563 %

Kernel Polinomial
-------------------

C Ordem do polinômio Erro de classificação (média dos folds)

0,25 2 22,6563 %

0,5 2 22,6563 %

1 2 22,9167 %

0,25 3 22,9167 %

0,5 3 23,0469 %

1 3 22,6563 %

0,25 4 23,9583 %

0,5 4 23,9583 %

1 4 23,8281 %

0,25 5 24,2188 %

0,5 5 24,2188 %

1 5 24,7396 %

Kernel de Base Radial
-----------------------

C gamma Erro de classificação (média dos folds)

0,25 0,001 34,8958 %

0,5 0,001 34,8958 %

1 0,001 34,8958 %

0,25 0,01 34,8958 %

0,5	0,01	34,8958 %
1	0,01	34,8958 %
0,25	0,1	34,8958 %
0,5	0,1	30,9896 %
1	0,1	24,0885 %
0,25	1	22,6563 %
0,5	1	22,6563 %
1	1	22,2656 %

Como se pode observar, para valores de  $C$  altos, a otimização da SVM escolhe um hiperplano de margem menor caso o hiperplano realize uma boa separação/classificação ao classificar todos os pontos de treinamento corretamente. Por outro lado, um valor de  $C$  muito pequeno fará com que o otimizador procure um hiperplano de separação de margem maior, mesmo que o mesmo classifica incorretamente mais pontos do conjunto de treinamento. Para o kernel Polinomial quanto maior a ordem, mais a margem pode se adequar ao espaço ocupado pelos dados, mas se a ordem do kernel for muito elevada, pode gerar overfitting.