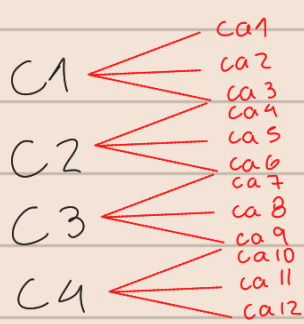


Classificação

Data Augmentation

ERRADO



② shuffling
+
separação
treino e teste

ca10 - C4
ca2 - C1
ca5 - C2
ca8 - C3
ca3 - C1
ca11 - C4
ca4 - C2
ca1 - C1
ca12 - C4
ca7 - C3
ca6 - C2
ca9 - C3

X_train desbalanceado
por ter C1 tanto
no treino quanto no teste

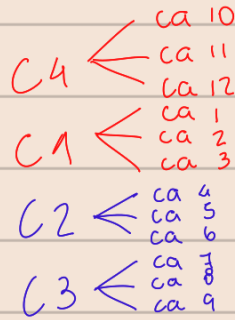
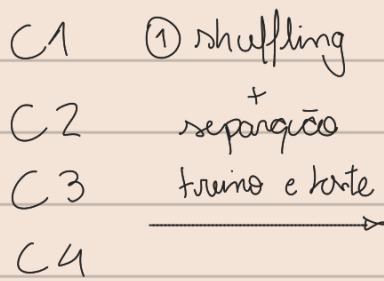
model fit (X_train, Y_train)

y_pred = model.predict(X_test)

acc = accuracy_score(y_pred, y_test)

① Data augmentation

CERTO



② Data augmentation

Matriz de confusão

diagonal é o que importa

↳ overfitting com uma diagonal muito perfeita

REAL		
	PREVISÃO	

precision → dos classificados positivos qnt % são realmente positivos

recall
sensibilidade → dos verdadeiramente positivos qnt % foram identificados como positivos

1 - Especificidade → dos negativos qnt % foram classificados como positivos

Sensibilidade + (1 - Especificidade) NÃO SOMAM 1

F1-score → combinação da média harmônica da precision e recall

$$F_1 = \frac{2}{\frac{1}{precision} + \frac{1}{recall}}$$

Validação cruzada - predict

CV = 3

dataset
treino

1	2	3
---	---	---

Experimento

Treina em

Prevê em

Predições honestas

1

1 + 2

3

y-pred honesto para 3

2

2 + 3

1

y-pred honesto para 2

3

3 + 1

2

y-pred honesto para 1