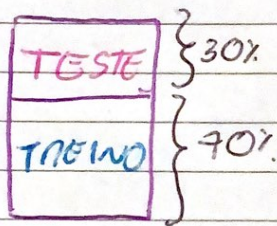


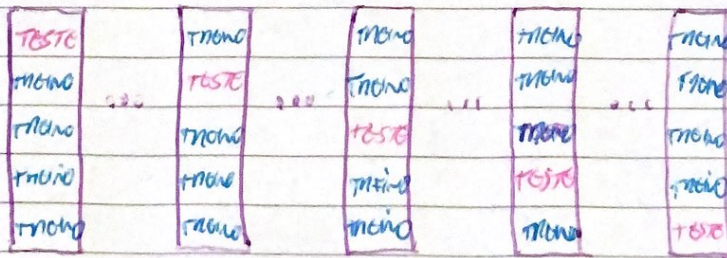
## VALIDAÇÃO CRUZADA K FOLD

ANteriormente, dividíamos nossos dados de treino e teste de forma aleatória, estipulando apenas qual seria a porcentagem de cada um. Como consequência, a pontuação (score) final variava de forma sucinta cada vez que executávamos o algoritmo.



ATRAVÉS DA VALIDAÇÃO CRUZADA, O CONJUNTO DE DADOS SERÁ DIVIDIDO EM K PARTES IGUAIS. NA PRIMEIRA ITERAÇÃO, UMA DAS PARTES SERÁ UTILIZADA PARA TESTE ENQUANTO AS DEMAS PARA TREINO. NAS PRÓXIMAS ITERAÇÕES EM DIANTE, OUTRAS PARTES QUE SE FOMAM USADAS PARA TESTE SERÃO USADAS PARA TREINO. PARA CADA ITERAÇÃO HAVENDO UM COEFICIENTE DE DETERMINAÇÃO  $R^2$  DIFERENTE, VEJO UM EXEMPLO

$$K = 5$$



$$0,73 + 0,71 + 0,75 + 0,79 + 0,70 = 0,936$$

5

VALORES RECOMENDADOS DE K: [5, 10]

OBS. MODELO UTILIZADO APENAS PARA ANÁLISE DE PERFORMANCE.