

Redes Neurais e Aprendizagem Profunda

APRENDIZADO DE MÁQUINA AJUSTE DE HIPERPARÂMETROS

Zenilton K. G. Patrocínio Jr

zenilton@pucminas.br

Recapitulando: Classificador Vizinho Mais Próximo

Como avaliar a similaridade entre imagens?  Uso de distância

Distância L1: $d_1(I_1, I_2) = \sum_p |I_1^p - I_2^p|$

Distância L2: $d_2(I_1, I_2) = \sqrt{\sum_p (I_1^p - I_2^p)^2}$

Exemplo com Distância L1

Imagen de teste				-	Imagen de treino				=	Valor abs. diferenças por pixel			
56	32	10	18		10	20	24	17		46	12	14	1
90	23	128	133		8	10	89	100		82	13	39	33
24	26	178	200		12	16	178	170		12	10	0	30
2	0	255	220		4	32	233	112		2	32	22	108

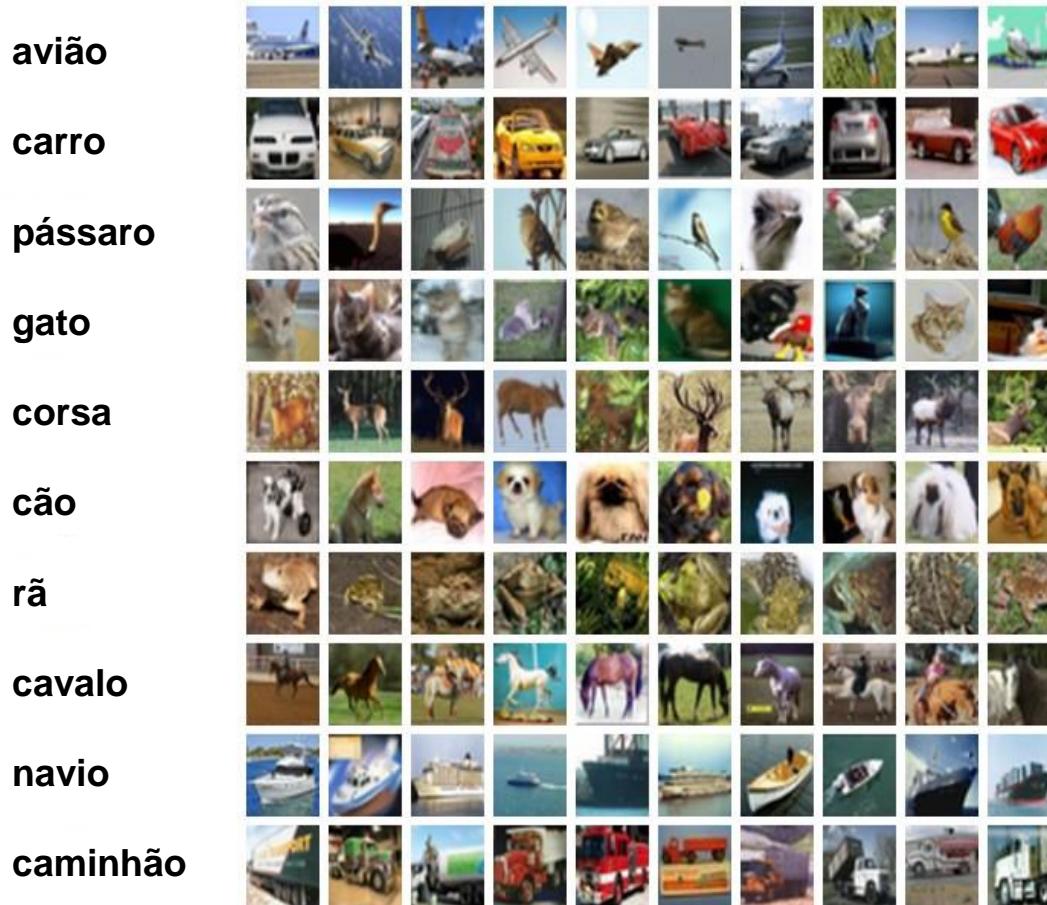
$\Sigma \rightarrow 456$

Recapitulando: Vizinho Mais Próximo – Dataset CIFAR-10

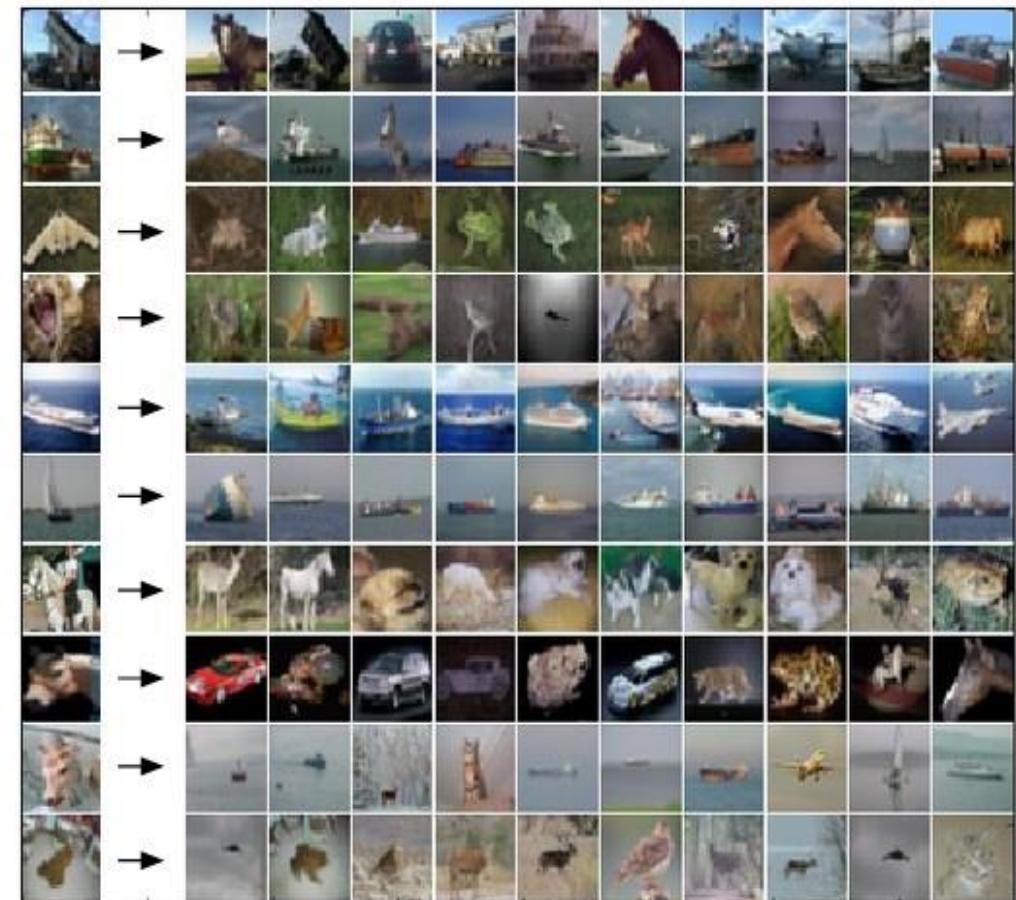
10 classes

50.000 imagens para treino

10.000 imagens para teste



Exemplos dos vizinhos mais próximos
para cada imagem teste (primeira coluna)



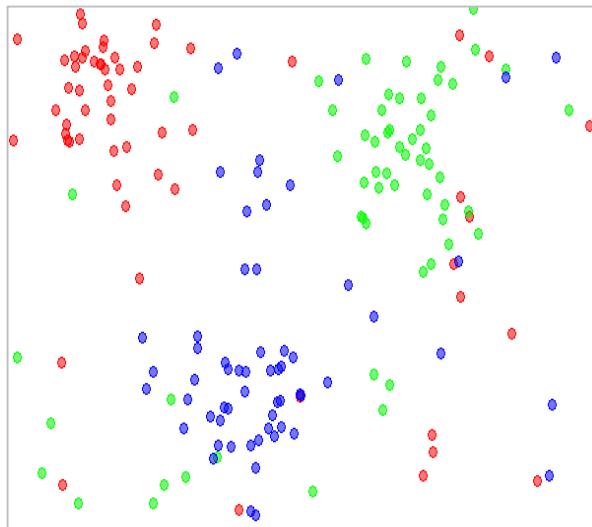
Classificador K Vizinhos Mais Próximos

Encontram-se as k amostras (imagens) mais próximas de uma nova observação e se realiza uma votação para se determinar seu rótulo

Classificador K Vizinhos Mais Próximos

Encontram-se as k amostras (imagens) mais próximas de uma nova observação e se realiza uma votação para se determinar seu rótulo

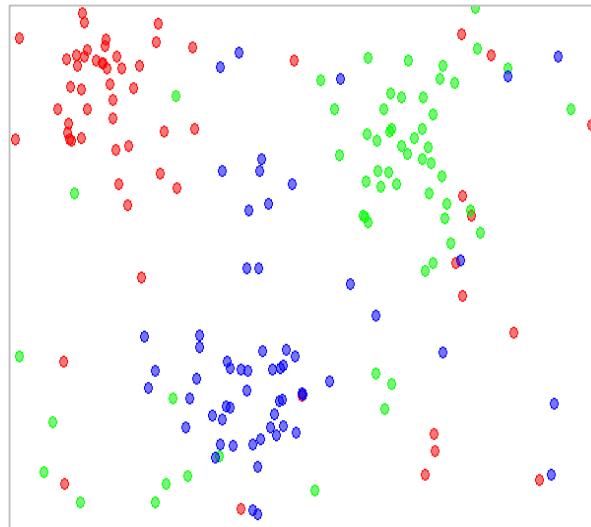
Dados



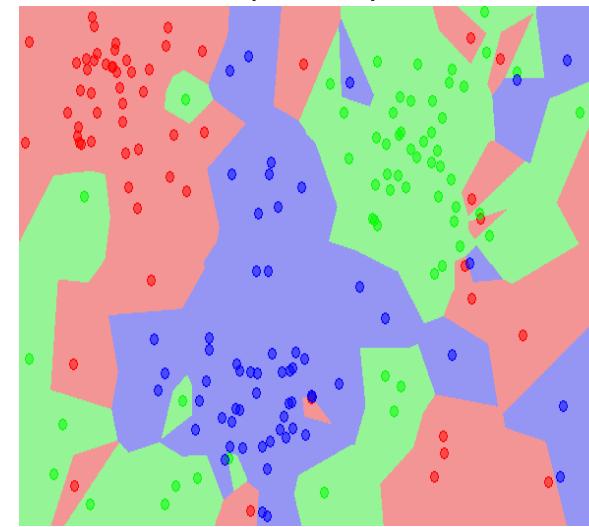
Classificador K Vizinhos Mais Próximos

Encontram-se as k amostras (imagens) mais próximas de uma nova observação e se realiza uma votação para se determinar seu rótulo

Dados



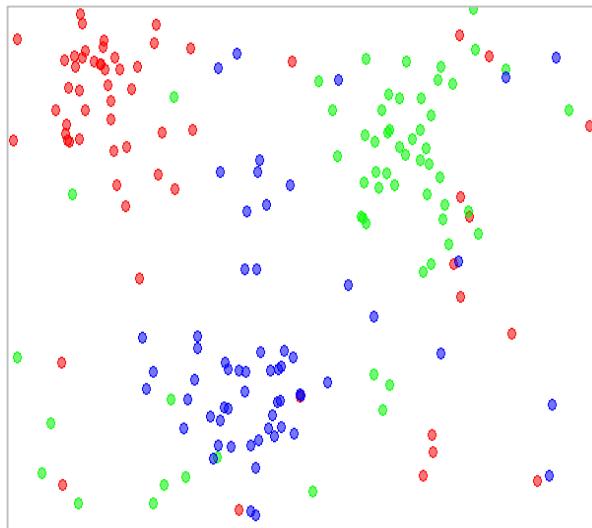
Classif. Vizinho + Próximo
(1-NN)



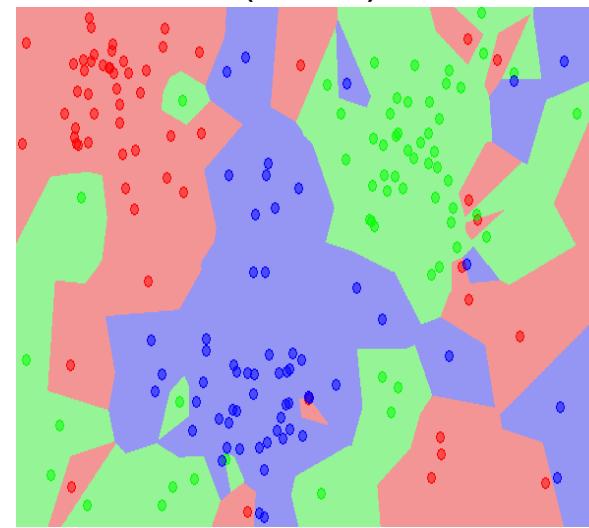
Classificador K Vizinhos Mais Próximos

Encontram-se as k amostras (imagens) mais próximas de uma nova observação e se realiza uma votação para se determinar seu rótulo

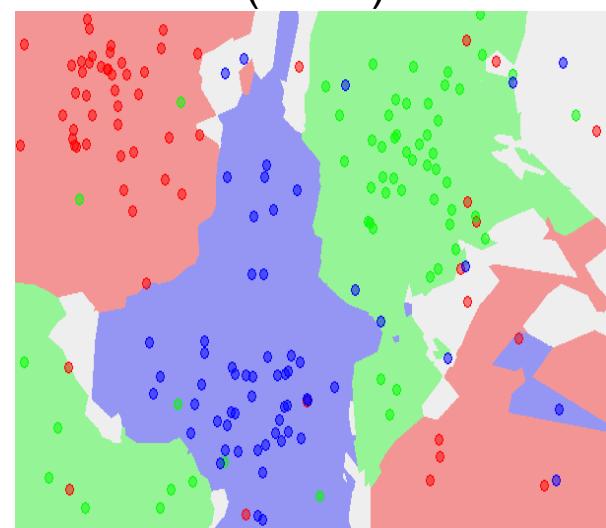
Dados



Classif. Vizinho + Próximo
(1-NN)

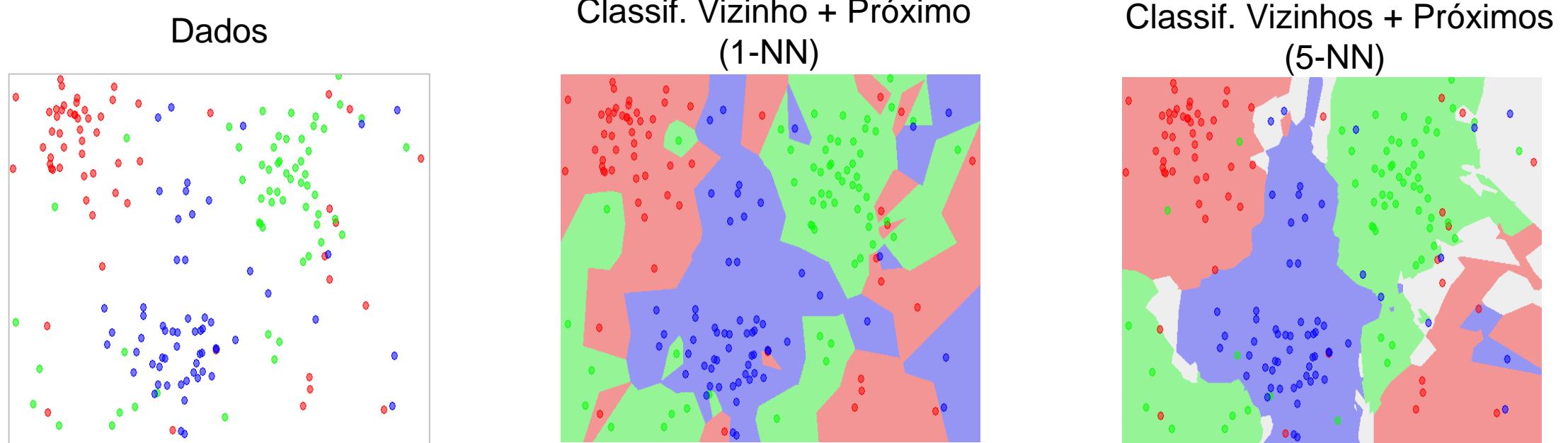


Classif. Vizinhos + Próximos
(5-NN)



Classificador K Vizinhos Mais Próximos

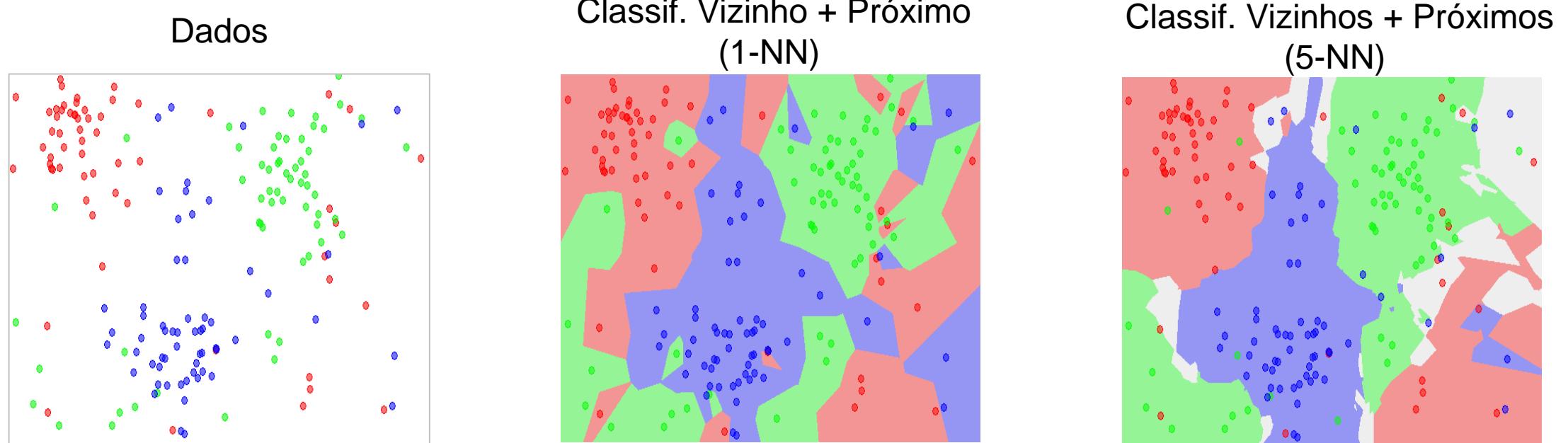
Encontram-se as k amostras (imagens) mais próximas de uma nova observação e se realiza uma votação para se determinar seu rótulo



P1: qual a acurácia do classificador sobre os dados de treinamento, quando se utilizar uma outra distância como, por exemplo, a Euclidiana?

Classificador K Vizinhos Mais Próximos

Encontram-se as k amostras (imagens) mais próximas de uma nova observação e se realiza uma votação para se determinar seu rótulo



P2: qual a acurácia do classificador sobre os dados de treinamento, quando se utilizar um número k de vizinhos diferente, por exemplo, 2 ou 5 ou 10?

Ajuste de Hiperparâmetros

Perguntas importantes ao se utilizar o classificador de vizinho(s) mais próximo(s)

- Qual é a melhor medida de **distância** para ser utilizada?
- Qual o número ideal de vizinhos **k** a ser usado?

Ajuste de Hiperparâmetros

Perguntas importantes ao se utilizar o classificador de vizinho(s) mais próximo(s)

- Qual é a melhor medida de **distância** para ser utilizada?
- Qual o número ideal de vizinhos **k** a ser usado?

Em outras palavras: como determinar estes **hiperparâmetros**?

Ajuste de Hiperparâmetros

Perguntas importantes ao se utilizar o classificador de vizinho(s) mais próximo(s)

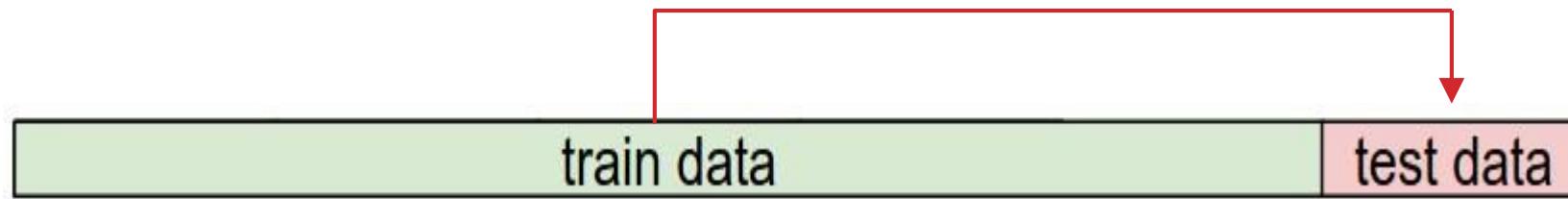
- Qual é a melhor medida de **distância** para ser utilizada?
- Qual o número ideal de vizinhos **k** a ser usado?

Em outras palavras: como determinar estes **hiperparâmetros**?

- Geralmente, eles são muito dependentes do problema
- Muitas vezes deve-se tentar “todas as combinações possíveis” e verificar qual produz o melhor resultado

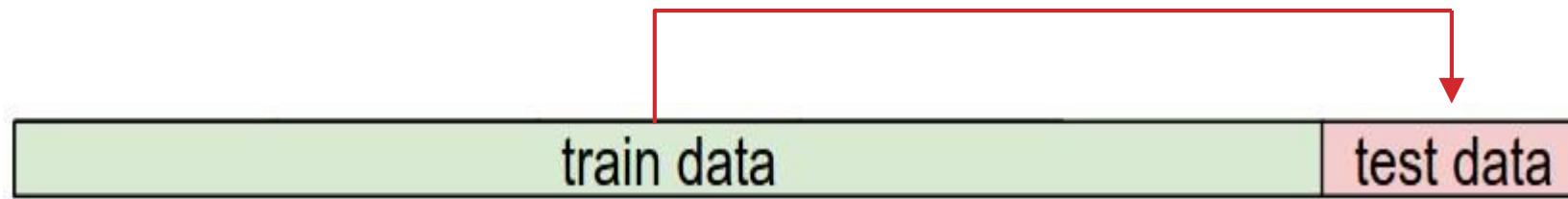
Ajuste de Hiperparâmetros

Opção 1: Experimentar quais hiperparâmetros funcionam melhor sobre o conjunto de teste



Ajuste de Hiperparâmetros

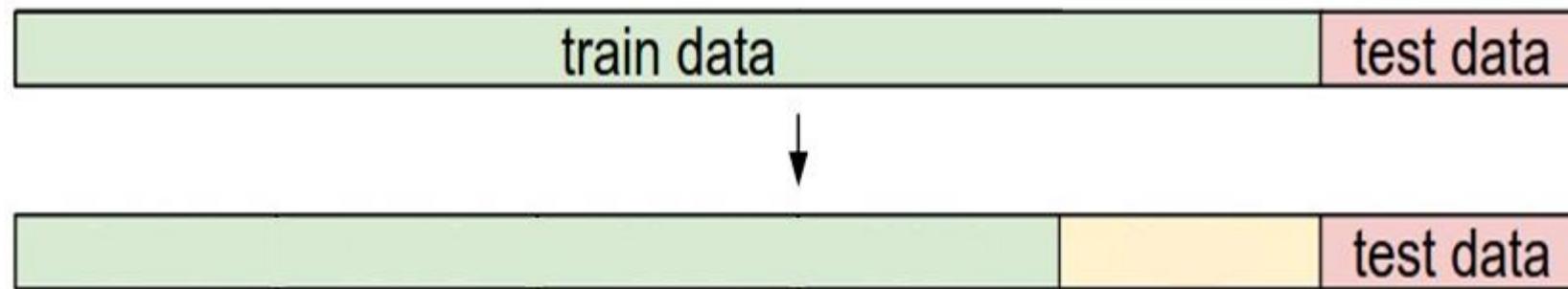
Opção 1: Experimentar quais hiperparâmetros funcionam melhor sobre o conjunto de teste



- Péssima ideia. O conjunto de teste deve ser utilizado para se avaliar a performance de generalização!
- Usa-se o conjunto de teste de forma **BEM ESPARSA**, no final de todo o processo

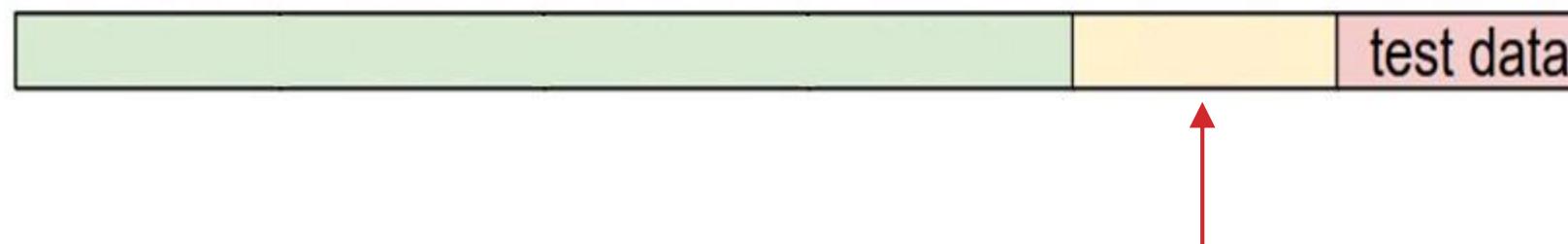
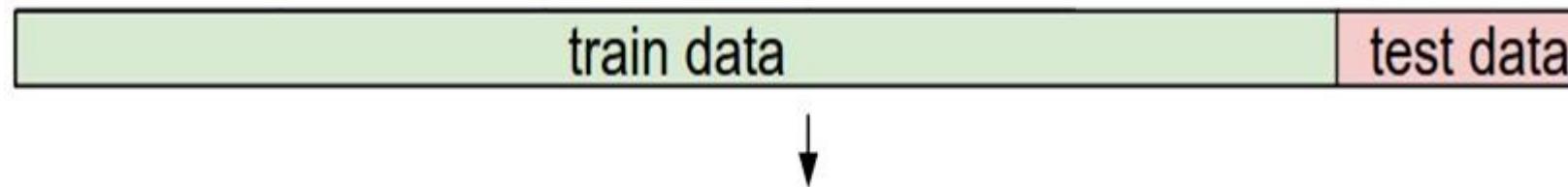
Ajuste de Hiperparâmetros

Opção 2: Dividir os dados de treinamento → criar um conjunto de validação



Ajuste de Hiperparâmetros

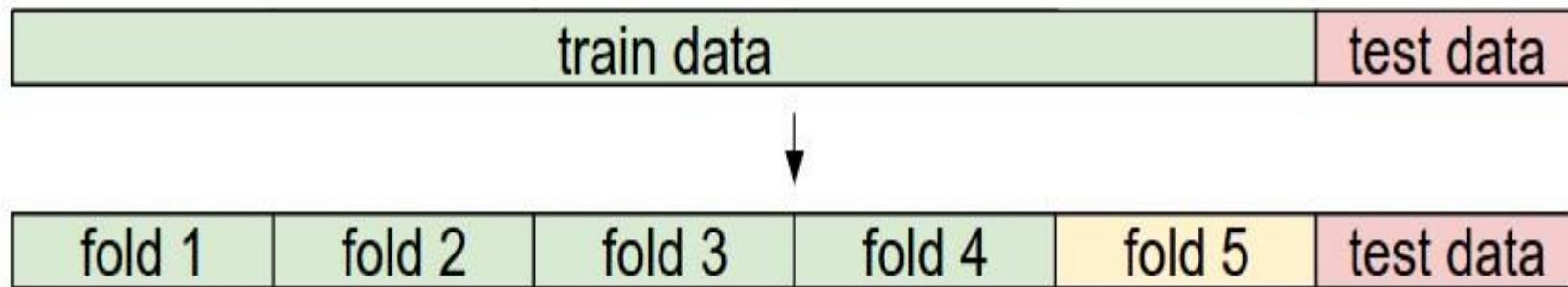
Opção 2: Dividir os dados de treinamento → criar um conjunto de validação



Dados de validação
usados para ajustar hiperparâmetros

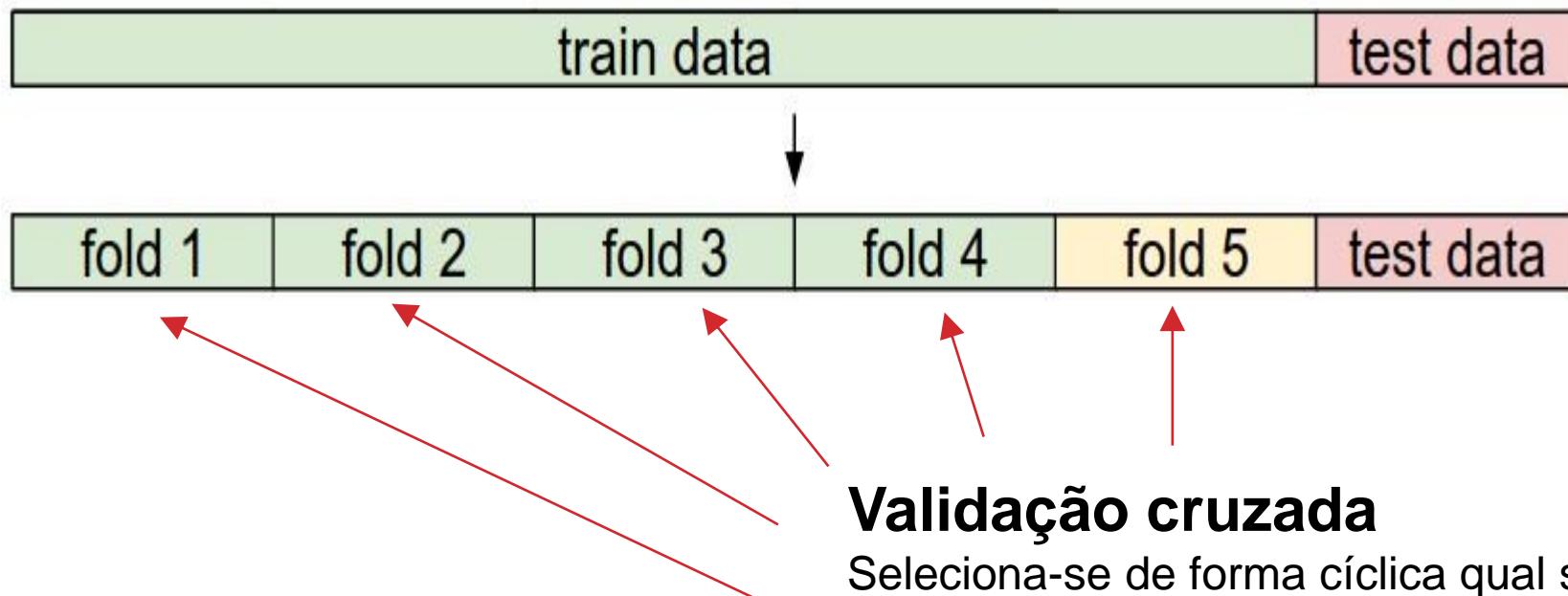
Ajuste de Hiperparâmetros

Opção 3: Dividir várias vezes os dados de treinamento → validação cruzada



Ajuste de Hiperparâmetros

Opção 3: Dividir várias vezes os dados de treinamento → validação cruzada

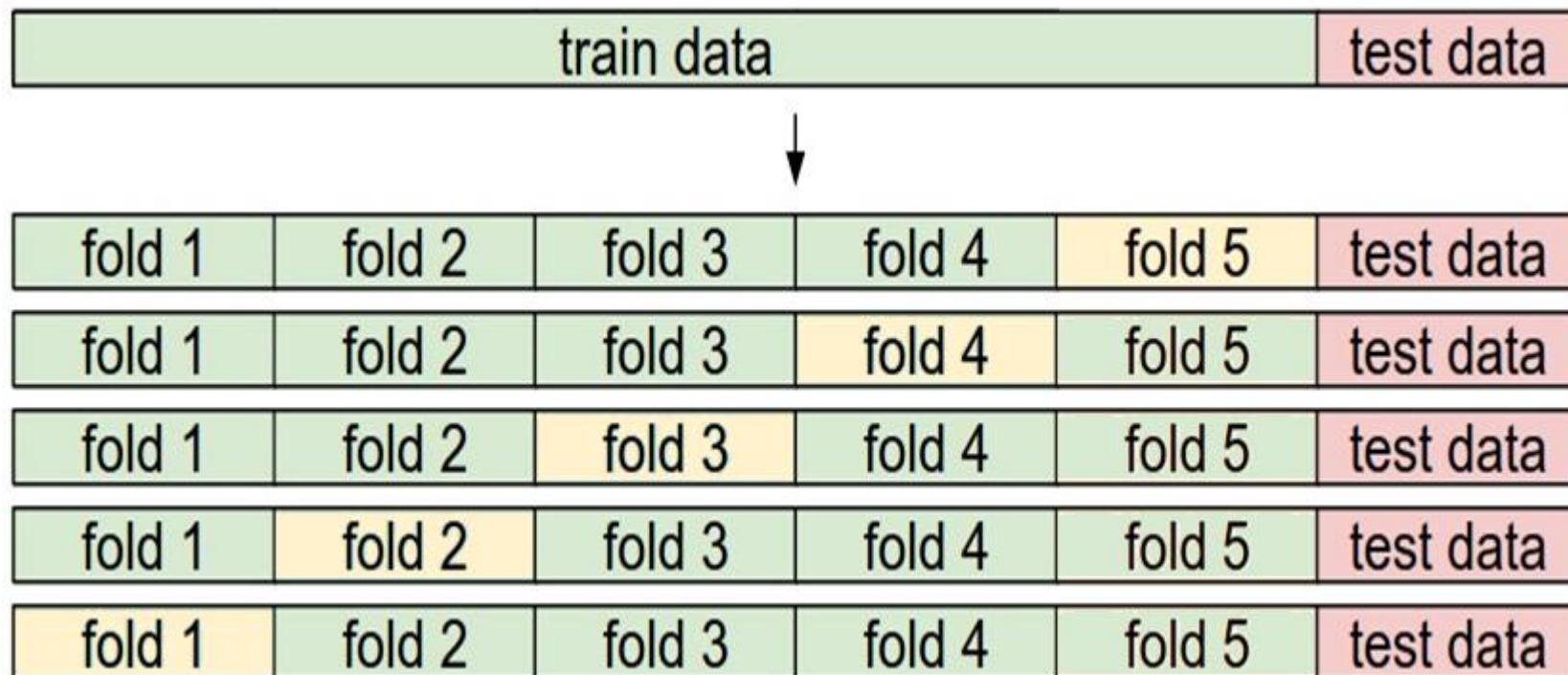


Validação cruzada

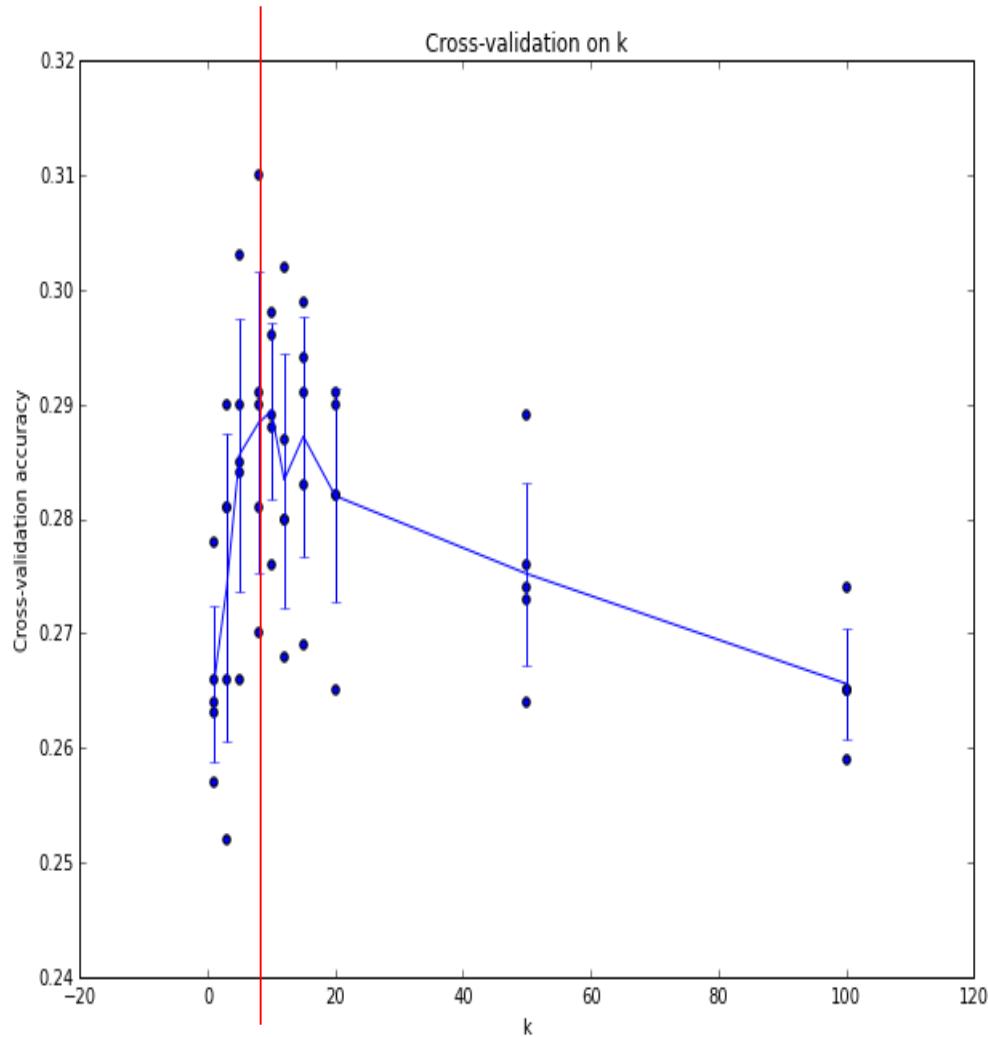
Seleciona-se de forma cíclica qual subconjunto (ou “*fold*”) será usado para validação e se toma a média dos resultados

Ajuste de Hiperparâmetros

Opção 3: Dividir várias vezes os dados de treinamento → validação cruzada

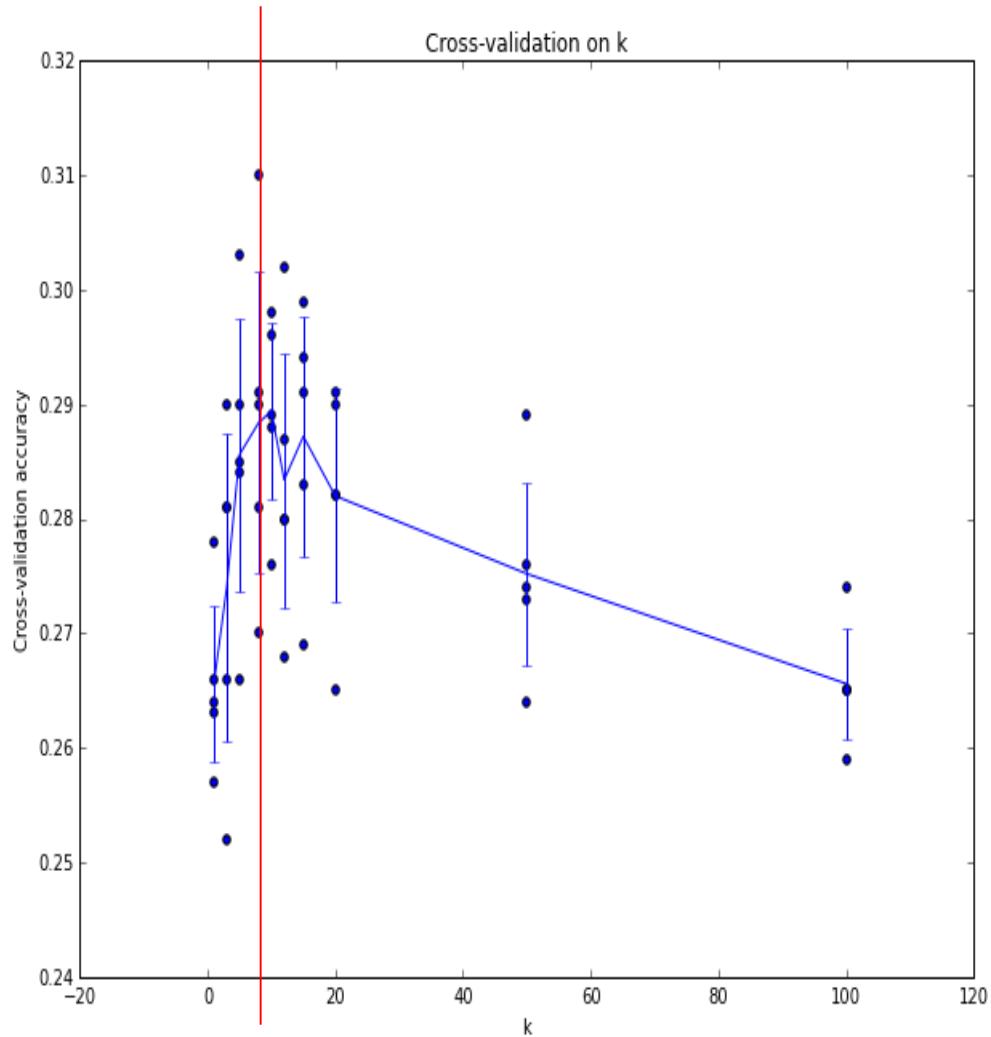


Exemplo – Ajuste de Hiperparâmetros



Exemplo de validação cruzada de 5 dobras (subconjuntos ou “folds”) para se determinar o valor de k

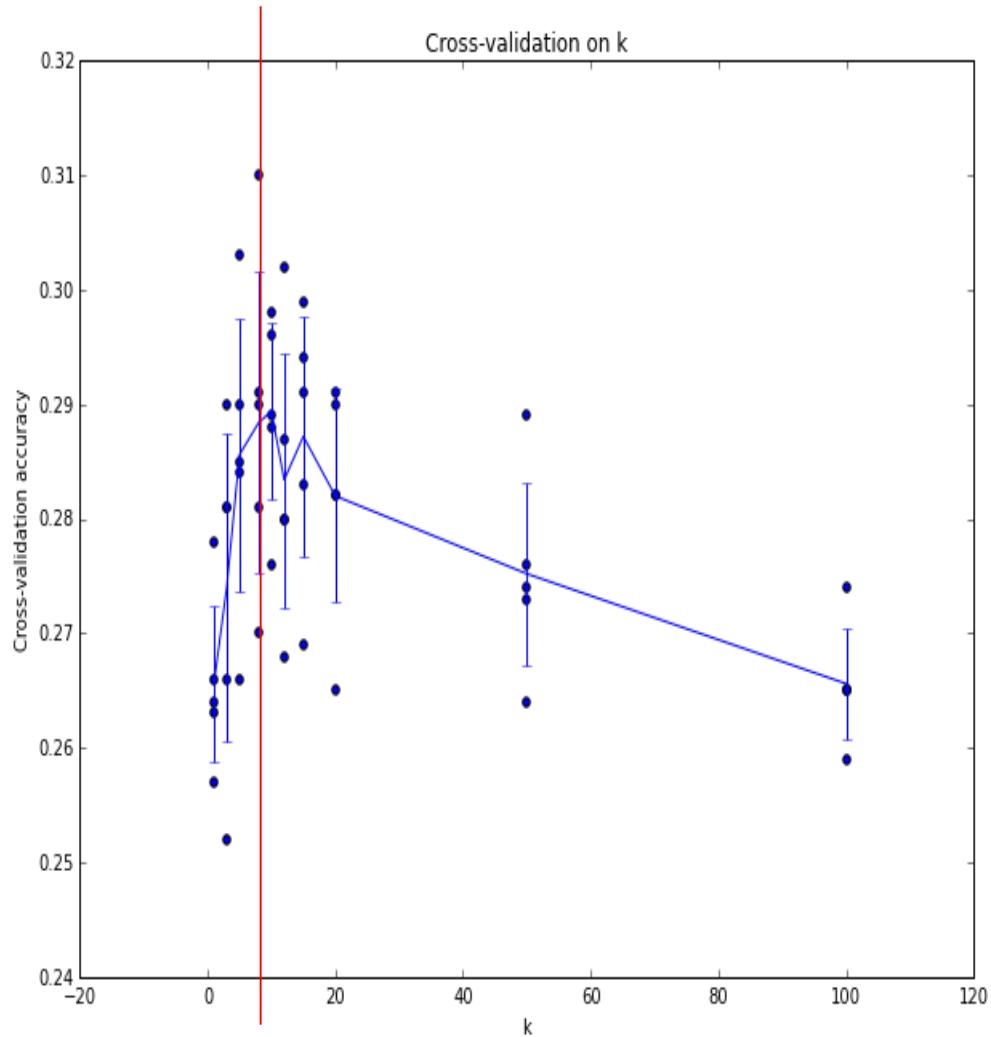
Exemplo – Ajuste de Hiperparâmetros



Exemplo de validação cruzada de 5 dobras (subconjuntos ou “folds”) para se determinar o valor de k

Cada ponto no gráfica representa um único resultado de acurácia

Exemplo – Ajuste de Hiperparâmetros

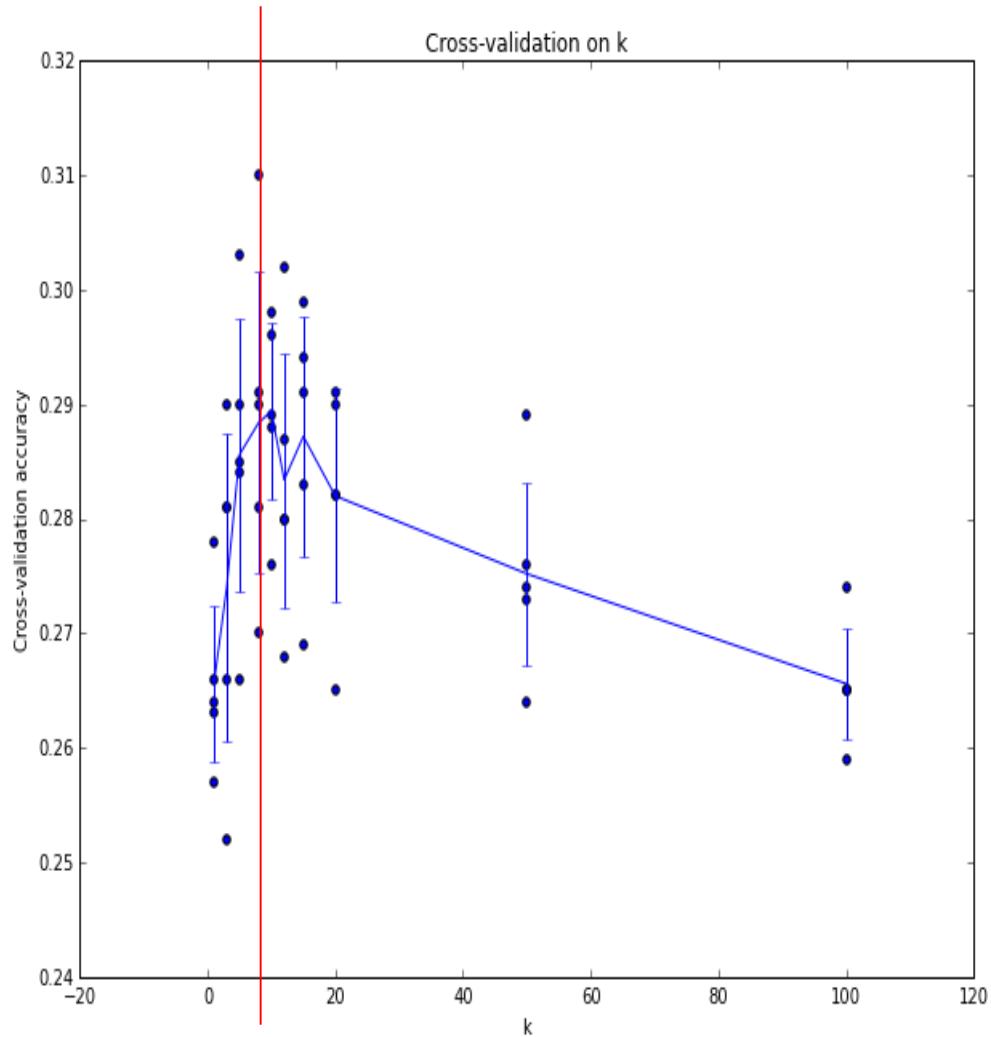


Exemplo de validação cruzada de 5 dobras (subconjuntos ou “folds”) para se determinar o valor de k

Cada ponto no gráfica representa um único resultado de acurácia

A linha azul liga as médias para diferentes valores de k , enquanto as barras indicam o desvio padrão

Exemplo – Ajuste de Hiperparâmetros



Exemplo de validação cruzada de 5 dobras (subconjuntos ou “folds”) para se determinar o valor de k

Cada ponto no gráfica representa um único resultado de acurácia

A linha azul liga as médias para diferentes valores de k , enquanto as barras indicam o desvio padrão

Parece que $k \approx 7$ apresenta a melhor média de resultados para esse conjunto de dados