

## Data Science com Python 3

### K-Vizinhos mais próximos

1 Adicione o atributo intenção de voto à base de dados referente à rede social de cientistas de dados. Sorteie uma intenção de voto para cada usuário da rede. Aumente o tamanho da base de modo que ela tenha 100 usuários.

2. Há uma técnica para validação de predições realizadas por algoritmos de aprendizado de máquina chamada ***cross validation leave one out***. Ela funciona da seguinte forma.

1. O algoritmo de predição é executado utilizando-se a base inteira exceto uma das instâncias.
2. O resultado obtido é comparado com o valor real.
3. O processo é repetido para cada uma das instâncias.
4. Ao final, exibe-se o percentual de acerto e erro.

Execute o algoritmo KNN (use  $K=5$ ) da forma descrita e verifique o resultado. Note que os dados foram gerados aleatoriamente e, portanto, usuários “próximos” não necessariamente têm intenção de voto igual.

### ***Bibliografia***

GRUS, J. **Data Science from Scratch: First Principles with Python**. 1st ed. O'Reilly Media, 2015.