

bootstrapping

February 10, 2019

```
In [1]: # https://en.wikipedia.org/wiki/Bootstrapping\_\(statistics\)
        # allows data repeat witch might be problematic

import random
def Bootstrap(n, n_iter=3, random_state=None):
    """
    Data generator for cross validation. In each iteration will generate set of indexes
    Returns data sample and list of all omitted indexes
    """
    if random_state:
        random.seed(random_state)
    for j in range(n_iter):
        bs = [random.randint(0, n-1) for i in range(n)]
        out_bs = list({i for i in range(n)} - set(bs))
        yield bs, out_bs

boot = Bootstrap(n=100, n_iter=10, random_state=1)
for train_idx, validation_idx in boot:
    print(train_idx, validation_idx)

[17, 72, 97, 8, 32, 15, 63, 97, 57, 60, 83, 48, 26, 12, 62, 3, 49, 55, 77, 97, 98, 0, 89, 57, 3,
[65, 13, 99, 20, 66, 50, 47, 62, 93, 3, 60, 5, 39, 90, 78, 75, 74, 50, 82, 21, 21, 64, 29, 1, 9,
[32, 67, 21, 84, 34, 82, 91, 37, 58, 89, 41, 63, 60, 14, 3, 39, 49, 43, 53, 24, 33, 13, 32, 93,
[85, 49, 37, 64, 63, 2, 41, 78, 51, 36, 2, 20, 25, 41, 72, 17, 43, 54, 27, 34, 86, 12, 48, 70,
[75, 53, 20, 14, 57, 21, 87, 30, 20, 95, 13, 55, 48, 69, 37, 70, 32, 91, 61, 40, 12, 26, 83, 40,
[96, 45, 63, 60, 19, 12, 64, 99, 41, 9, 65, 85, 22, 22, 99, 19, 18, 40, 39, 13, 90, 65, 77, 37,
[40, 63, 87, 61, 28, 91, 52, 43, 71, 78, 93, 83, 35, 82, 28, 6, 9, 97, 65, 82, 47, 20, 65, 98,
[31, 48, 95, 71, 0, 24, 67, 56, 74, 2, 3, 80, 77, 31, 33, 26, 22, 36, 18, 69, 25, 34, 39, 74, 9,
[39, 89, 21, 57, 79, 85, 67, 25, 46, 67, 0, 86, 49, 74, 54, 51, 43, 79, 74, 93, 89, 95, 8, 63,
[6, 21, 38, 83, 94, 91, 71, 34, 45, 78, 94, 29, 50, 71, 51, 22, 61, 33, 78, 42, 91, 28, 33, 78]
```

```
In [ ]:
```