

K-MEANS

1. 对python中 count(), values_counts(), size() 函数:

① str.count(sub, start=0, end=len(string))

② list.count(obj) → 统计变量obj在list中的次数

③ pandas中 values_counts() 函数对series每个值进行计数并排序。比如 df['区域'].value_counts() 默认降序。value_counts() 返回series数组

1) value_counts(normalize=True) 归一化

2) ascending=True 改为升序。

0 shape 和 size 既可作为函数, 也可作为ndarray属性。

2. random.choices(population, weights=None, *, cum_weights=None, k=1) 种群 相对权重 每次取数

3. k-means 的 n_init 参数, 用于设置聚类中心初始值及迭代次数。最终解在merit意义下达到最优结果, init参数若指定 random, 则随机从训练数据中选取初始值。

ax = Axes3D(fig, rect=[0, 0, .95, 1], elev=48, azim=134) 将函数 仰角, 方位角, 默认30。

optional 图形, 又图形, 方位角初角, 默认为60。 (left, bottom, width, height) 轴位置。

ax.w_xaxis.set_ticklabels(12) 显示刻度。

4. 创建文本: text=ax.text(2)(x, y, text)

bbox是text3D的一个参数, bbox=dict(alpha=.2, edgecolor='w', facecolor='w') 透明度

设置矩形的颜色及字体设置框

比如
5. `np.choose(y, [1, 2, 0]).astype(np.float)`
`np.choose(a, choices, out=None, mode='raise')`
按 `a` 对 `choices` 选择。 `a` 中数 $[0, n-1]$ 范围

比如: `a = [1, 0, 1]`

`choices = [10, 10]`

`np.choose(a, choices) → array([10, 10, 10])`
映射

6. sklearn 的 `metrics` 模块实现了一些 loss, score 以及一些工具函数计算分类性能。

`silhouette_score(X, labels, metric='euclidean')`
轮廓系数 欧氏距离

7. `sklearn.datasets.make_blobs()` 是用于创建多类样本数据集函数, 为每个类分配一个或多个正态分布的总集。

8. 随机数是由随机种子根据一定的算法计算出来的数值。

只要算法一定, 随机种子一定, 产生随机数就变。计算机产生的伪随机数既是随机, 又是有规律的。

设置随机数种子, 为了保证每次随机数一致。