1，cluster数目>100,作为数据集

result 0.9035087719298246, 0.916208346813329, 0.927109028120375, 0.9284892086330935, 0.9249158539742816, 0.9489840034587116, 0.9519203003176436, 0.9567954841511073, 0.956140350877193

2，做其他三个数据集

Biomedical

未预处理

0.5206666666666667, 0.5859375, 0.6244285714285714, 0.6491666666666667, 0.6713, 0.691, 0.7218333333333333, 0.74925, 0.791

[0.5383888888888889, 0.5771875, 0.6234285714285714, 0.6455833333333333, 0.6658, 0.68275, 0.7085, 0.734, 0.769]

**[0.6842247433819557, 0.7760332252836305, 0.8165084510303311, 0.8253646677471637, 0.8427876823338736, 0.8375202593192869, 0.8592652620205294, 0.8496758508914101, 0.8565640194489466]**

[0.7566666666666667, 0.7894375, 0.8039285714285714, 0.8203333333333334, 0.8295, 0.8345, 0.8416666666666667, 0.8525, 0.857]

Nb

[0.55105555555555552, 0.57956249999999998, 0.61071428571428577, 0.65266666666666662, 0.66590000000000005, 0.68687500000000001, 0.70699999999999996, 0.73450000000000004, 0.76849999999999996]

[0.5280929227444624, 0.59187601296596437, 0.69425793007640657, 0.72730956239870337, 0.77066450567260936, 0.78525121555915722, 0.79119394921663966, 0.79902755267423009, 0.80307941653160453]

[0.81938888888888894, 0.83787500000000004, 0.84464285714285714, 0.84850000000000003, 0.85099999999999998, 0.85187500000000005, 0.85350000000000004, 0.86024999999999996, 0.86250000000000004]

[0.57694444444444448, 0.60293750000000002, 0.62564285714285717, 0.66283333333333339, 0.67530000000000001, 0.70287500000000003, 0.71516666666666662, 0.75249999999999995, 0.79249999999999998]

[0.58148748424275165, 0.66136547811993518, 0.74820560314887707, 0.78133441383036195, 0.82106969205834679, 0.8320502431118314, 0.83495407887628315, 0.83711507293354948, 0.83468395461912481]

[0.83483333333333332, 0.85668750000000005, 0.86742857142857144, 0.87450000000000006, 0.87460000000000004, 0.87912500000000005, 0.88066666666666671, 0.88100000000000001, 0.88549999999999995]

[[0.42429292929292928, 0.44691919191919194]]

[[0.32683805469133781, 0.38038316685770429]]

[[0.7606565656565657, 0.77060606060606063]]

[0.36297979797979796]

[0.05182577370230883]

[0.6915151515151515]

[0.3793434343434343]

[0.4017520877681349]

[0.6861616161616162]

[0.21348484848484847]

**[0.5596037334206648]**

[0.5238383838383839]

[0.17525252525252524]

[0.5683641722613394]

[0.3842929292929293]

[0.258989898989899]

[0.4954150974291796]

[0.5091414141414141]

1

试一下beta，迭代次数不同时的效果；

2

词典的大小V是动态的；

3

self.n\_z[cluster] += wordFre \* prob

4

高概率document用于更新model