Week 11

TARGET:

- find implemention of BP and more ML techniques in python.
- calculate more characteristics of the brains.
- store them in a decent way.

ML techniques

Found a repo of some ML techniques implemented in python including BP, SVM, K-Means, etc.

https://github.com/lawlite19/MachineLearning_Python

I tried BP. The accuracy is pretty low. Just a little bit better than guessing.

预测准确度为： 35.294118%

Thats what reminds me that "K_Means is not funtional in the project" is not true. Because the earlier demos are limited to 3 dimensions in which no ML method can perform corrcetly.

The urgent need is acquire more characteristics but bot blame ML methods

More characteristics

chars that have been calculated:

- assortativity
- charpath
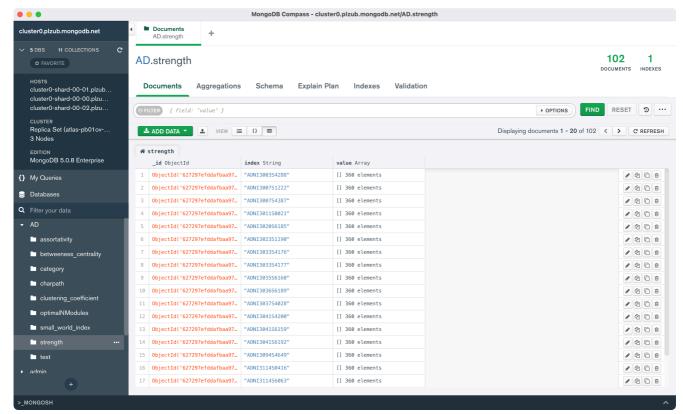- small_world_index
- mean clustering_coefficient

Done this week

- clustering_coefficient (360 vector)
- maximized modularity: 将所有点分组，使得组内节点连接最大化，组间最小化。output: [Ci, Q]. Q is the maximized modularity.
- mean clustering coefficient: clustering coefficient 的平均值。
- sthength (360 vector)
- local efficiency (360 vector)
- betweenness centrality (360 vector)

manage the data

1. store matlab varables in json.

⊗ ⊘ 3.LMCI.json   ↥ Open with Xcode

[{"index":"ADNI3002S4229","value":
[2344.2631578947326,249.52631578947424,352.68421052631368,2588.4736842105203,1215.8421052631543,2095.8421052631597,194.789473684211
,1796.8947368421013,377.94736842104845,1068.4736842105242,255.84210526315852,464.26315789473125,390.57894736841797,398.999999999999
03,24.263157894737027,95.842105263158416,702.15789473683924,1266.3684210526292,1055.8421052631588,609.52631578947148,855.8421052631
6042.8506.368421052639,55.842105263158174,163.21052631578991,1150.5789473684181,24.26315789473696,1222.1578947368384,645.3157894736
8462,870.57894736842013,2133.7368421052597,167.42105263157941,895.84210526315769,30.578947368421254,17.947368421052627,17.947368421
052627,1449.5263157894701,205.3157894736476,2457.947368421052,68.473684210526727,601.10526315789093,582.1578947368389,430.57894736
841894,47.421052631579386,17.947368421052627,847.42105263157839,346.36842105262923,1057.9473684210509,1969.5263157894674,83.2105263
15789892,3274.78947368421,447.42105263157322,1998.999999999993,3398.9999999999986,22.157894736842231,535.84210526315633,1106.368421
0526278,17.947368421052627,0,51.631578947368787,17.947368421052627,0,57.947368421052957,887.42105263157578,17.947368421052627,445.3
1578947368428,17.947368421052627,544.26315789473267,1645.3157894736805,190.57894736842155,203.21052631579,228.47368421052681,723.21
052631578675,2291.6315789473624,415.84210526315695,756.89473684210088,1215.8421052631518,1697.9473684210463,3253.7368421052638,346.
36842105262991,17.947368421052627,1841.1052631578902,1184.2631578947335,683.21052631579,1348.473684210514,1182.1578947368384,3788.4
736842105262,706.3684210526319,62.157894736842422,777.94736842104885,97.947368421053,17.947368421052627,17.947368421052627,455.8421
052631578,285.31578947368439,17.947368421052627,2192.6842105263104,64.263157894737333,295.84210526315712,1615.8421052631531,1691.63
15789473626,788.47368421051476,26.368421052631785,1601.1052631578871,3786.368421052633,20.05263157894743,430.57894736841678,17.9473
68421052627,464.26315789473233,68.47368421052667,17.947368421052627,17.947368421052627,17.947368421052627,41.105263157895052,17.947
368421052627,872.68421052631379,1744.2631578947307,1582.1578947368353,0,0,986.36842105263054,17.947368421052627,159.0000000000043,
291.63157894736747,312.68421052631209,1436.8947368421098,17.947368421052627,85.31578947368449,990.5789473684182,1072.6842105263136,
1893.7368421052597,0,201.10526315789519,95.842105263158317,1022.1578947368348,441.10526315789076,182.15789473684262,2933.7368421052
615,472.6842105263126,161.10526315789525,1293.7368421052568,184.26315789473739,775.84210526315849,636.89473684210407,2365.315789473
6823,3525.3157894736814,784.26315789473642,420.05263157894365,1316.8947368420963,1236.8947368421016,2645.3157894736883,215.84210526
31584,4013.7368421052652,66.36842105263193,493.73684210526267,17.947368421052627,28.473684210526325,203.21052631579002,1266.3684210
52628,708.47368421052306,17.947368421052627,17.947368421052627,22.15789473684222,91.63157894736865,424.26315789473722,17.947368421
052627,85.31578947368 4589,424.26315789473352,1740.0526315789407,512.68421052631356,910.57894736841,47.421052631579272,17.9473684210
52627,0,401.10526315789053,1824.2631578947776,474.78947368420557,860.05263157894615,733.73684210525266,453.73684210525943,17.947368
421052627,2535.8421052631543,43.210526315789643,266.36842105263167,1165.3157894736846,497.94736842105237,1072.6842105263154,215.842
10526315846,1196.894736842102,1329.5263157894715,329.52631578947188,17.947368421052627,314.78947368421086,137.947368421053,813.7368
4210526392,161.10526315789519,74.789473684210989,85.31578947368474,4653.7368421052633,100.05263157894763,476.894736842102,2224.263
1578947344,990.57894736841956,81.1052631578949,22.157894736842135,594.78947368420972,17.947368421052627,1466.3684210526262,420.0526
3157894689,708.473684210525,287.42105263157896,26.368421052631746,864.26315789473631,17.947368421052627,601.10526315789571,60.05263
1578947725,1074.7894736842095,1702.1578947368362,575.84210526315655,100.05263157894775,525.315789473682,272.68421052631567,558.9999
9999999977,255.84210526315837,41.105263157894996,453.73684210526096,502.15789473684038,388.4736842105246,582.15789473684049,272.684
21052631578,603.21052631578857,420.05263157894603,2297.947368421052,1969.526315789471,45.315789473684305,32.68421052636015,921.105
26315789332,17.947368421052627,24.263157894736967,1121.1052631578928,373.73684210526244,241.10526315789522,87.421052631579272,270.5
7894736842081,220.05263157894788,17.947368421052627,17.947368421052627,472.68421052631419,137.94736842105306,990.57894736841945,17.
947368421052627,428.473684210525,234.78947368421095,1285.3157894736792,100.05263157894758,211.63157894736887,17.947368421052627,287
.42105263157833,636.89473684210475,0,17.947368421052627,257.94736842105289,714.78947368421166,49.526315789473905,748.47368421052659
,603.21052631578834,3491.6315789473679,2411.6315789473724,17.947368421052627,588.47368421052659,17.947368421052627,1537.94736842105
02,17.947368421052627,32.68421052631595,17.947368421052627,55.84210526315146,832.68421052631709,68.473684210526571,26.36842105263
1721,337.94736842105112,413.73684210526267,1815.842105263155,215.84210526315843,0,186.36842105263193,57.9473684210528,179,17.947368
421052627,51.63157894736856,17.947368421052627,0,733.73684210526051,17.947368421052627,0,0,17.947368421052627,295.84210526315786,55
.842105263158089,0,26.368421052631604,17.947368421052627,0,17.947368421052627,432.6842105263151,2276.8947368421045,17.94736842105
2627,255.84210526315834,817.94736842105021,104.26315789473705,420.05263157894723,17.947368421052627,140.05263157894771,1457.9473684
210507,925.31578947368268,306.36842105263226,17.947368421052627,752.68421052636315,17.94736842105217,69.34090909008468,1484.7954545
052829,1079.0000000000009,2430.578947368419,53.736842105263335,76.8947368421056,582.15789473684174,1870.5789473684181,784.263157894
7362,422.15789473684134,624.26315789473665,1403.2105263157875,407.42105263157788,57.947368421052822,1447.4210526315737,28.473684210
526525,0,135.84210526315837,60.05263157894576,171.63157894736643,2105263157888,2154.7894736842027,17.947368421052627,62.157894
736842323,60.052631578947683,228.47368421052673,34.789473684210762,660.05263157894524,179,45.315789473684404,161.10526315789519,161
7.9473684210509,17.947368421052627,24.263157894736867,0,331.63157894736628,420.05263157894711,1592.6842105263149,241.10526315789502
,17.947368421052627,119.00000000000037,17.947368421052627]},{"index":"ADNI3002S4654","value":
[2530.0227272727379,40.704545454545674,136.84090909090872,993.88636363635976,302.52272727272543,916.15909090908792,14.1136363636364
81,560.24999999999864,226.84090909091015,1503.2045454545403,36.613636363636537,934.56818181817721,128.65909090908994,417.0681818181
792,672.74999999999682,239.11363636363768,601.15909090908987,2096.3863636363621,67.29545454545418,222.75000000000114,527.52272727272
646,791.3863636363609,16.15909090909102,7.9772727272727009,455.93181818181472,61.159090909091127,3186.61363636336,7.97727272727270
09,580.70454545454231,1895.931818181811,1261.8409090909038,7.9772727272727009,46.8409090909093,7.9772727272727009,1286.386363636362
8,7.9772727272727009,1895.9318181818137,284.1136363636362,1286.3863636363587,1073.6590909090867,128.65909090908929,12.0681818181819
2,1026.6136363636315,128.65909090908954,981.61363636363285,691.15909090908792,1499.1136363636315,725.93181818181461,83.659090909090
239,948.88636363636158,1728.0454545454421,406.84090909090907,1507.2954545454502,110.24999999999872,69.34090909090468,1484.7954545454
495,453.88636363636209,136.84090909090847,87.749999999999517,723.88636363636112,7.9772727272727009,48.88636363636389,1357.97727272
7268,1400.9318181818142,7.9772727272727009,218.65909090909184,854.79545454545075,474.34090909090577,1462.2954545454497,343.43181818
1818,14.113636363636438,640.022727272724,2045.2499999999944,437.52272727272458,744.34090909090571,120.47727272727141,719.7954545454
5416,5011.1590909090755,16.159090909091006,110.24999999999875,361.84090909090514,1865.2499999999964,772.97727272726968,3374.7954545
4546,7.9772727272727009,963.20454545454186,1566.6136363636322,138.88636363636473,7.9772727272727009,204.3409090909083,7.97727272727
27009,7.9772727272727009,723.88636363635544,1889.7954545454393,181.84090909090963,1685.2499999999936,1687.2954545454493,192.0681818
181819,7.9772727272727009,2984.1136363636429,766.84090909090673,519.34090909090742,30.477272727272883,625.70454545454186,973.431818
18181506,367.97727272726939,715.70454545454174,1587.068181818177,247.29545454545558,171.61363636363564,1045.0227272727257,63.204545
454545624,159.34090909090875,365.93181818181846,173.659090909092,151.15909090909025,783.2045454545414,0,255.47727272727363,318.8863
63636362,183.88636363636468,496.84090909090884,7.9772727272727009,592.97727272726934,7.9772727272727009,7.9772727272727009,498.8863

2. read the json files into mongoDB.

3. combine characteristics into one csv file in Python.

```
code  >  田 test.csv
     1    index,category,assortativity,charpath,small_world_index
     2    ADNI3003S4288,HC,0.320225447,3.0050979117287797,1.60916555
     3    ADNI3011S0021,HC,0.361080855,3.6270242198624874,1.01825523
     4    ADNI3941S6094,HC,0.242998973,4.774371986292598,0.596358538
     5    ADNI3036S6189,HC,0.323415786,4.032268742967868,0.807971358
     6    ADNI3007S4387,HC,0.0927308351,2.609213754752773,1.84740877
     7    ADNI3168S6131,HC,0.287759215,4.0968601704261305,0.925716639
     8    ADNI3041S4200,HC,0.30425182,3.922824385917449,0.93542695
     9    ADNI3041S6159,HC,0.361986101,3.6950910136394315,1.19517922
    10    ADNI3168S6085,HC,0.326990157,3.9347148434466135,0.97074616
    11    ADNI3137S4482,HC,0.129289493,2.824344238910419,1.50160706
    12    ADNI3037S4028,HC,0.336414546,3.9431628587835847,0.793561578
    13    ADNI3020S6185,HC,0.44435057,3.5623507306851963,1.12135041
    14    ADNI3007S1222,HC,0.132521719,3.837387902197926,0.91436106
    15    ADNI3114S6063,HC,0.0820066929,2.7168343024534503,1.63931394
    16    ADNI3114S0416,HC,0.150920436,3.252531356077069,1.32921875
    17    ADNI3033S4176,HC,0.0937950537,3.2477696290854734,1.35910058
    18    ADNI3168S6098,HC,0.218100414,4.220405872929106,0.927514255
    19    ADNI3041S6192,HC,0.400519907,4.318266945890552,0.806505
    20    ADNI3033S4177,HC,0.412268102,4.040450832092194,1.22737968
    21    ADNI3116S6119,HC,0.36996457,5.038633036090718,0.720885873
    22    ADNI3094S4649,HC,0.272836506,3.45943129239136,1.29266536
    23    ADNI3116S4453,HC,0.40680787,4.058885601237838,0.909498
    24    ADNI3035S6160,HC,0.294244409,3.4056059639360767,1.13217306
    25    ADNI3023S1190,HC,0.456693709,3.194368750994278,1.81653023
    26    ADNI3021S4659,EMCI,0.171734482,2.168953375213609,2.74835443
    27    ADNI3002S4473,EMCI,0.366116554,4.795743209587524,0.744117439
    28    ADNI3037S4706,EMCI,0.264353275,4.06752796039204,0.96108371
```

4. perform ML methods on these csv files.

One more thing

After calculating a little, I've got **1445** characteristics. input them to the BP network.

```
>>> a.shape
(102, 1446)
```

BP network 4 classes

预测准确度为：83.333333%

K-Means 2 classes

$$0.722772277227227$$

Note that raw data is not divided into training set and testing set. But the Correct Rate is still better than last week.

TODO

- divide the raw data into training set and testing set.
- see what will happen then.

PCA