The performance VS. Feature Addition

We have implemented the algorithm "Feature Addition" mentioned in the paper "Dynamic Feature Selection with Fuzzy-Rough Sets" in Matlab. In paper "Dynamic Feature Selection with Fuzzy-Rough Sets", it introduced two similarity relations:

1) $u_{R_a}(x,y) = 1 - \dfrac{|a(x) - a(y)|}{a_{max} - a_{min}}$ 　　　　2) $u_{R_a}(x,y) = \exp(-\dfrac{(a(x) - a(y))^2}{2\delta_a^2})$.

We have used these two relations to implement two algorithms named "Feature_Addition_1" and "Feature_Addition_2". The comparison performances of OFS-Density VS. Feature_Addition are as follows:

Table1: Predictive accuracy using KNN as basic classifier

Data set	OFS-Density	Feature_Addition_1	Feature_Addition_2
IONOSPHERE	0. 8771	0. 8686	0. 8686
WDBC	0. 9367	0. 9104	0. 9367
SONAR	0. 6288	0. 8318	0. 8657
HILL	0. 6149	1	1
COLON	0. 7833	0. 6333	0. 6333
SRBCT	0. 8167	0. 8333	0. 8667
LUNG2	0. 88	0. 925	0. 91
LYMPHOMA	0. 9667	0. 95	0. 9167
GLIOMA	0. 72	0. 52	0. 72
MLL	0. 9	0. 9571	0. 7
PROSTATE	0. 87	0. 67	0. 7
DLBCL	0. 825	0. 8875	0. 6475
LEU	0. 9286	0. 7143	0. 7286
ARCENE	0. 85	0. 81	0. 62
AVG.	0. 8284	0. 8222	0. 7938

Table2: Predictive accuracy using SVM as basic classifier

Data set	OFS-Density	Feature_Addition_1	Feature_Addition_2
IONOSPHERE	0. 82	0. 92	0. 92
WDBC	0. 9508	0. 9385	0. 9701
SONAR	0. 7023	0. 7737	0. 7341
HILL	0. 5339	1	1
COLON	0. 8167	0. 7833	0. 75
SRBCT	0. 8333	0. 95	0. 9667
LUNG2	0. 88	0. 92	0. 915
LYMPHOMA	0. 9833	0. 9833	0. 9833
GLIOMA	0. 64	0. 7	0. 7
MLL	0. 8571	0. 9714	0. 7429
PROSTATE	0. 94	0. 89	0. 86

DLBCL	0. 8125	0. 975	0. 68
LEU	0. 9286	0. 9571	0. 9571
ARCENE	0. 815	0. 805	0. 675
AVG.	0. 8223	0. 8976	0. 8467

Table3: Predictive accuracy using SVM as basic classifier

Data set	OFS-Density	Feature_Addition_1	Feature_Addition_2
IONOSPHERE	0. 8686	0. 8743	0. 8743
WDBC	0. 9227	0. 921	0. 9456
SONAR	0. 6479	0. 7033	0. 6912
HILL	0. 6058	1	1
COLON	0. 75	0. 7	0. 6667
SRBCT	0. 8333	0. 5667	0. 7333
LUNG2	0. 82	0. 89	0. 795
LYMPHOMA	0. 95	0. 7	0. 7167
GLIOMA	0. 64	0. 44	0. 4
MLL	0. 8143	0. 8	0. 6286
PROSTATE	0. 88	0. 75	0. 66
DLBCL	0. 8	0. 7875	0. 5725
LEU	0. 9143	0. 7	0. 7
ARCENE	0. 745	0. 745	0. 72
AVG.	0. 7994	0. 7555	0. 7217

Table4: The mean number of selected features

Data set	OFS-Density	Feature_Addition_1	Feature_Addition_2
IONOSPHERE	3. 6	34	34
WDBC	5. 4	7	30
SONAR	3. 8	43. 3	60
HILL	5. 8	83. 9	92. 2
COLON	5. 8	311. 4	551. 5
SRBCT	3. 2	372. 5	647. 2
LUNG2	13. 3	206. 8	280. 9
LYMPHOMA	15. 7	536. 7	733. 7
GLIOMA	5. 1	142. 9	54. 2
MLL	7. 8	162. 7	26
PROSTATE	5. 6	610	1230. 3
DLBCL	6. 7	157. 8	13. 7
LEU	4	534. 7	935. 1
ARCENE	51. 4	378. 5	21. 1
AVG.	9. 8	255. 8	336. 4

Table5: Running time (seconds)

Data set	OFS-Density	Feature_Addition_1	Feature_Addition_2

IONOSPHERE	0. 3129	1. 8659	0. 9397
WDBC	0. 5734	1. 4436	2. 1191
SONAR	0. 176	1. 9234	1. 0679
HILL	11. 1783	160. 7499	102. 4749
COLON	0. 9097	58. 3422	46. 641
SRBCT	0. 9728	86. 4838	64. 4031
LUNG2	7. 9992	571. 1522	37. 8869
LYMPHOMA	2. 2302	202. 9086	129. 0613
GLIOMA	1. 5975	45. 9449	0. 2558
MLL	2. 9385	127. 7084	0. 0493
PROSTATE	4. 4942	811. 7319	710. 4457
DLBCL	3. 5213	139. 3828	0. 0117
LEU	3. 6137	496. 6097	380. 1404
ARCENE	30. 7334	2871. 4018	0. 1632
AVG.	5. 08	398. 40	105. 40

In order to validate whether OFS-Density and Feature_Addition have significant difference in the prediction accuracy, we have conducted the Friedman test at a 95% significance level under the null-hypothesis. The p-values on KNN, SVM and CART are 0.6122, 0.2046 and 0.0957 respectively. Thus, there is no significant difference between OFS-Density and Feature_Addition on the prediction accuracy. However, form experimental results in Table 1 to Table 5, we can see that:

1) For some data sets, such as IONOSPHERE, WDBC, SONAR and HILL, Feature_Addition almost selected all the features. Main reason is that, the dependency of selected subset always cannot be up to 1 for real-world data sets. Thus, most features in these data sets increase the dependency, which makes the final selected subset contain almost all the features.

2) The running time of Feature_Addition is almost 20-80 times more than that in OFS-Density. As the paper says, "a fuzzy-rough evaluator may become less favorable due to its higher computational complexity".

In general, Feature_Addition is comparable to other online feature selection methods on the prediction accuracy. However, online streaming feature selection requires the efficiency, and Feature_Addition is unsuitable for streaming feature selection in online manner.