we provide several experimental results to have a comprehensive comparison of existing streaming feature selection algorithms, fast-OSFS [1], Alpha-investing [2] and Grafting [3]. Comparative algorithms compared were performed using their original implementations and settings. All experimental results are obtained on a personal computer with Windows 10, Inter(R) Core (TM) i3-4170 CPU (3.70 GHz) and 8.00 GB memory employing MATLAB R2015a platform. The details of benchmark data sets used is shown in Table 1. All data sets are high dimensional data from medical or healthy field.

TABLE 1. DETAILS OF MEDICAL DATA SETS

No.	Dataset	Instances	Features	Source
1	prostate	102	12600	mldataorg
2	central-nervous-sys	60	7129	mldataorg
3	lung-cancer-michigan	96	7129	mldataorg
4	leu	38	7129	libsvm
5	marti0	500	1024	ChaLearm
6	reged0	500	999	ChaLearm
7	arcene	100	10000	NIPS2003
8	madelon	2000	500	NIPS2003

Experiment 1 Results on the end of algorithms

Considering data sets with their natural order, results are collected using 10-fold validation at the end of algorithms. Features are handled one by one to simulate the scenario of streaming features. Figure 1 gives the size of selected features and Figure 2 shows the predictive accuracy by KNN classifier (k=3) using selected features. Figure 3 gives the average runtimes in 10-fold validation.

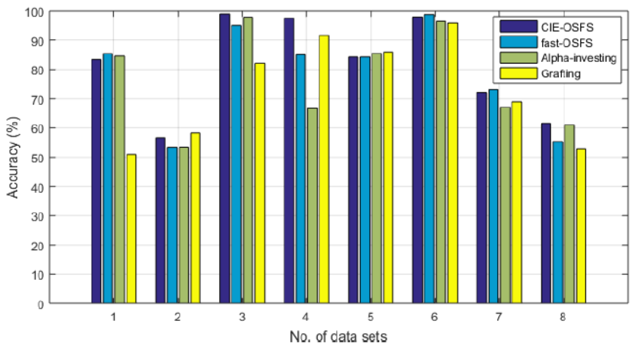

Figure 1. The predictive accuracy of algorithms on data sets

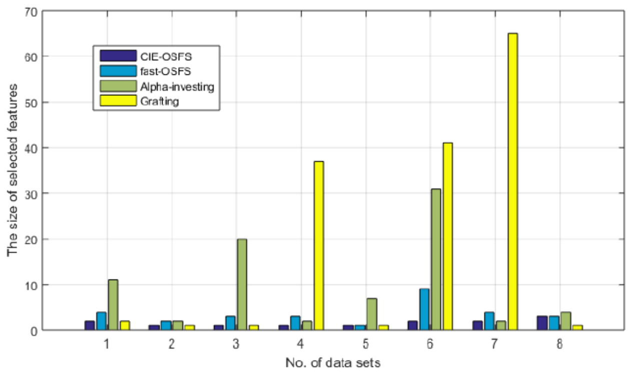

Figure 2. The size of selected features

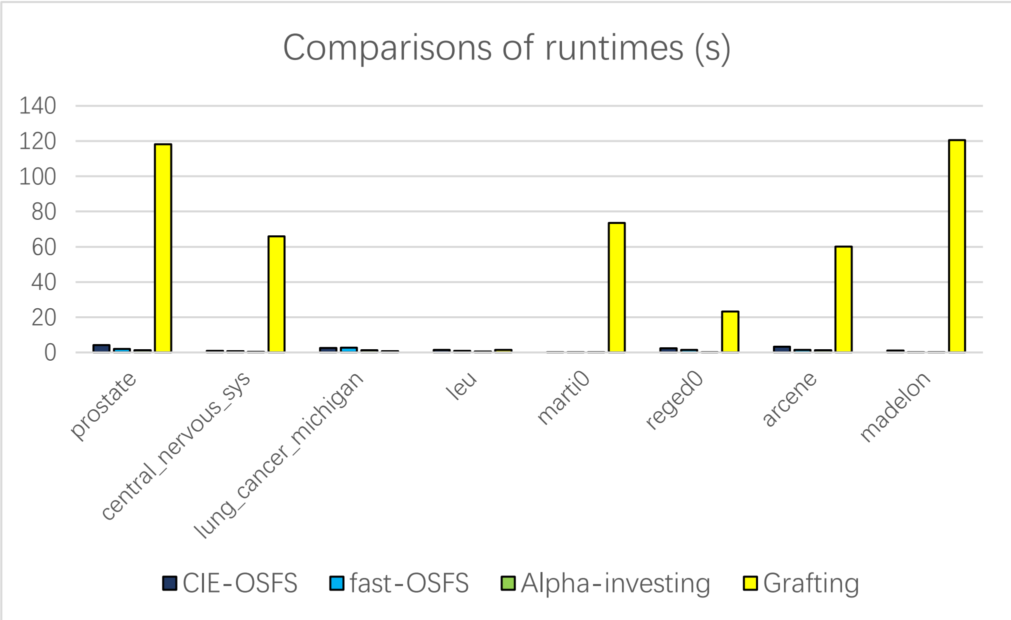

Figure 3. Comparisons of runtimes

Experiment 2 Comparisons of Order Stability

In real applications, features may arrive in random order. A good OSFS algorithm should be also robust to the diversification of the order of features. To estimate the stability of OSFS algorithms, we carry out experiments for ten times in different random orders. Figure 4 shows the mean accuracy of ten trails on 8 data sets. and Table 2 gives details of their selected subset.

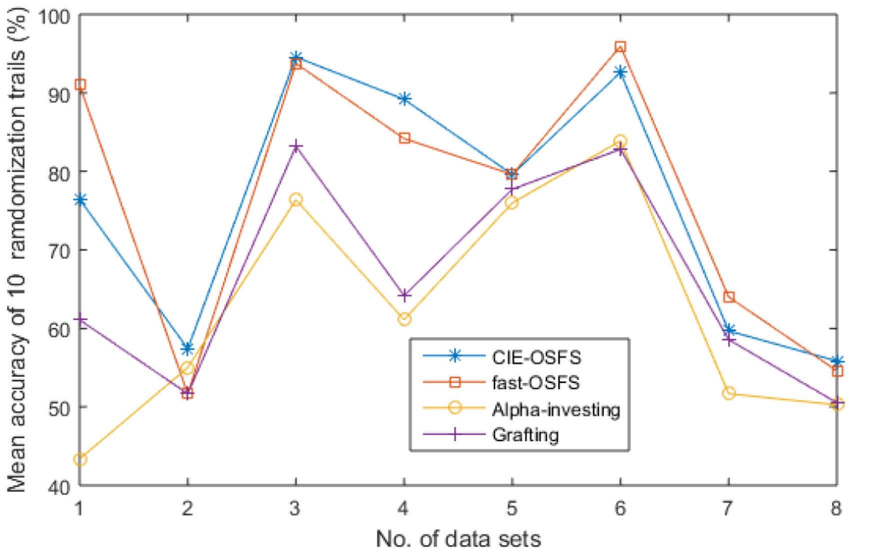

Figure 4. Comparisons of accuracy stability while the order of features changing

TABLE 2. Comparison of selected subset stability while the order of features changing

No.	CIE-OSFS			fast-OSFS			Alpha-investing			Grafting		
	min	max	mean	min	max	mean	min	max	mean	min	max	mean
1	**2**	**2**	**2**	4	6	5.4	8	14	10.1	**2**	**2**	**2**
2	**1**	2	1.2	2	4	3	1	5	2.6	**1**	**1**	**1**
3	**1**	**1**	**1**	3	5	3.9	15	28	21.5	**1**	**1**	**1**
4	**1**	**1**	**1**	3	5	3.8	**1**	5	2.3	39	47	43.2
5	**1**	**1**	**1**	**1**	**1**	**1**	6	41	24.2	**1**	**1**	**1**
6	**2**	**2**	**2**	10	12	11.2	30	43	36.3	37	74	51.3
7	2	**3**	**2.4**	4	5	4.3	**1**	5	2.1	61	71	67.6
8	3	**3**	**3**	**2**	5	3.6	4	7	5.4	3	338	34.7

Experiment 3 Simulating the OSFS scenario

Algorithms only need to assure the reliability of the output at the end of them in traditional feature selection problems. While getting any information about the entire feature space is impossible but the reliability of any temporary result during the process is required in the scenario of streaming features. In medical field, for example, tests patients will take is not accessible but we should ensure that the diagnosis from present evidences is correct at any time. To simulate the scenario of streaming feature selection and evaluate the effectiveness in process of OSFS algorithm, we adopt prediction accuracy and size of selected features to evaluate the process of streaming feature selection. In experiments, the first 10 percentage of features were handled at beginning and the rest of features were processed increasingly. Figure 5 gives the performance of four algorithms on 8 data sets.

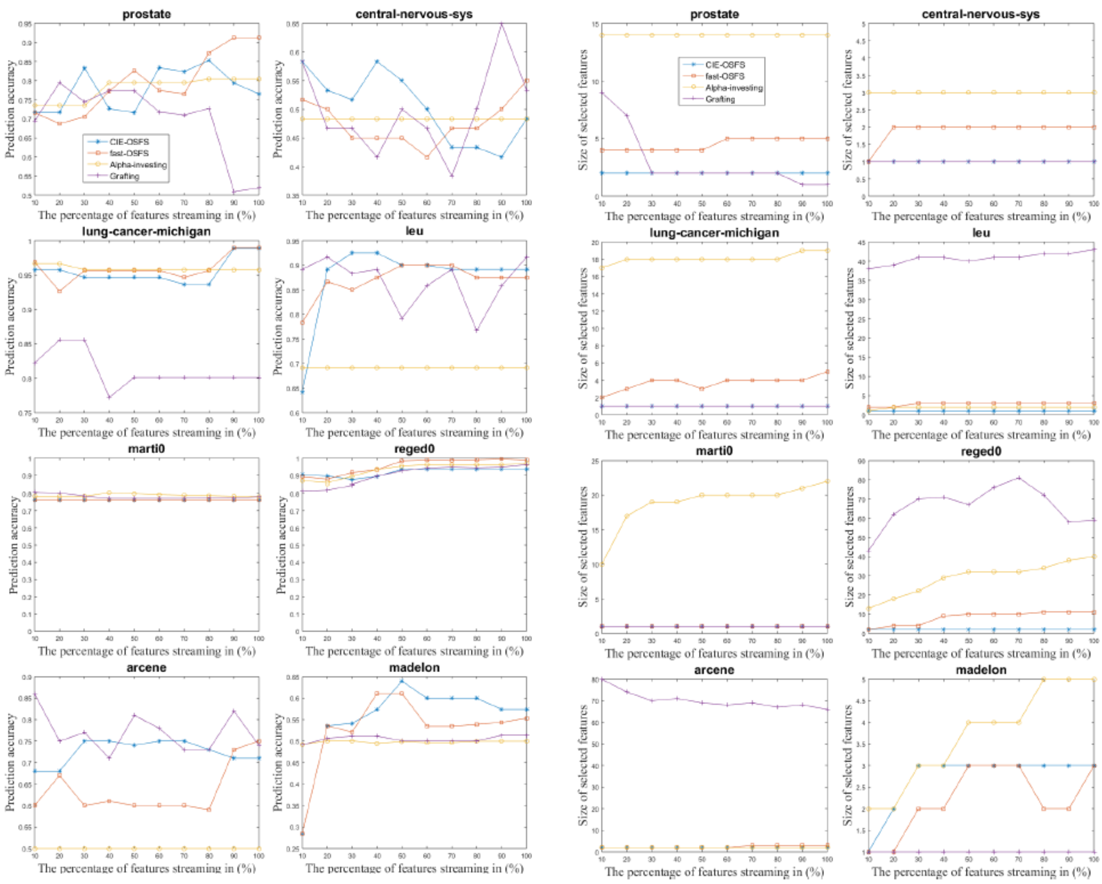

Figure 5. The performance of four algorithms in OSFS scenario

Reference

[1]. X. Wu, K. Yu, W. Ding, H. Wang, and X. Zhu, “Online feature selection with streaming features.” IEEE Transactions on Pattern Analysis & Machine Intelligence, vol. 35, no. 5, pp. 1178–1192, 2013.

[2]. J. Zhou, D. Foster, R. Stine, and L. Ungar, “Streaming feature selection using alpha-investing,” in Eleventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Chicago, Illinois, Usa, August, 2005, pp. 384–393.

[3]. S. Perkins and J. Theiler, “Online feature selection using grafting,” in Machine Learning, Proceedings of the Twentieth International Conference (ICML 2003), August 21-24, 2003, Washington, DC, USA, 2003, pp. 592–599.