Date: 02/10/18

1 Tasks

- To get lot of insights and classification do's and don'ts from 3 open datasets of Indian Pines ,Salinas and Pavia.
- To implement the classifiers kNN , RF , SVM , GMM , Adaboost , Ensemble
- To do data preprocessing like normalization , linear and non-linear dimensionality reduction
- Parameter tuning using Gridsearch and cross validation
- Feature selection by analysing mutual information matrix
- testing all the estimators/models and find OA , AA , Kappa , ROC , F1 score in each case

2 Classification Tables of Indian Pines

Table reading key -

- SVM-l refers to Linear kernel SVM
- SVM-p refers to Polynomial kernel SVM
- SVM-r refers to Radial Basis Function kernel SVM
- RF refers to Random Forest Classifier
- neigh refers to the number of neighbours used in kNN algorithm

Table 1: Indian Pines Dataset with train size 2%

Class	**Train**	**test**	**SVM-l**	**SVM-p**	**SVM-r**	**RF**	**kNN**
Alfalfa	0	46	0	0	0	0	0
Corn-notill	28	1400	45	53	0	70	41
Corn-mintill	16	814	45	15	0	0	38
Corn-mintill	4	233	39	21	0	0	41
Grass-pasture	9	474	51	48	0	73	44
Grass-trees	14	716	72	70	0	48	59
Grass-pasture-mowed	0	28	0	0	0	0	0
Hay-windrowed	9	469	82	83	0	72	80
Oats	0	20	0	0	0	0	0
Soybean-notill	19	953	47	53	0	0	30
Soybean-mintill	49	2406	48	41	24	38	51
Soybean-clean	11	582	20	37	0	0	13
Wheat	4	201	70	73	0	0	63
Woods	25	1240	82	82	0	77	76
Blds-grass-drives	7	379	40	45	0	0	0
Stone-steel-towers	1	92	0	0	0	0	0
OA			55.73	51.308	23.93	48.3736	51.885
AA			52	50	6	39	46
Kappa			0.4803	0.4169	0	0.3684	0.4389
Parameter grid			C=0.05 G=0.01	C=0.01 G=0.05	C=0.01 G=0.01	n_estimators 700 max_features log2	

3 Classification Tables of Salinas Dataset

Table 2: Indian Pines Dataset with train size=10 samples

Class	**Train**	**test**	**SVM-l**	**SVM-p**	**SVM-r**	**RF**	**kNN**	**GMM**
Alfalfa	10	36	18	16	14	15	9	0
Corn-notill	10	1418	68	60	29	21	43	5
Corn-mintill	10	820	30	37	39	12	23	0
Corn-mintill	10	227	29	25	18	0	9	0
Grass-pasture	10	473	52	64	61	39	39	0
Grass-trees	10	720	87	86	77	67	70	0
Grass-pasture-mowed	10	18	39	2	27	27	7	0
Hay-windrowed	10	468	96	93	91	61	92	0
Oats	10	10	5	3	3	3	2	0
Soybean-notill	10	962	35	36	31	0	29	0
Soybean-mintill	10	2445	62	72	58	0	57	0
Soybean-clean	10	583	26	36	15	0	17	0
Wheat	10	195	51	77	73	53	56	0
Woods	10	1255	87	89	85	75	82	0
Blds-grass-drives	10	376	58	42	19	54	15	0
Stone-steel-towers	10	83	75	77	100	39	99	0
OA			54.19	51.2439	41.659	34.0271	38.2495	0
AA			60	63	52	26	50	0.01
Kappa			0.4831	0.4572	0.3417	0.2605	0.3133	0.03
Parameter grid			C=0.05	C=0.02	C=0.01	n_est 700	neigh=8	comp=9
			G=0.01	G=0.05	Gamma=0.05	f=log2		cov=full

Table 3: Indian Pines Dataset with train size 10%

Class	**Train**	**test**	**SVM-l**	**SVM-p**	**SVM-r**	**RF**	**kNN**
Alfalfa	4	42	34	36	0	0	0
Corn-notill	142	1286	67	62	62	52	48
Corn-mintill	83	747	68	68	0	0	59
Corn-mintill	23	214	58	29	0	0	32
Grass-pasture	48	435	81	86	0	0	81
Grass-trees	73	657	85	80	52	34	70
Grass-pasture-mowed	2	26	100	0	0	0	0
Hay-windrowed	47	431	88	88	0	0	83
Oats	2	18	0	0	0	0	0
Soybean-notill	97	875	65	74	0	0	48
Soybean-mintill	245	2210	61	57	36	42	57
Soybean-clean	59	534	55	45	0	0	41
Wheat	20	185	94	91	0	0	77
Woods	126	1139	88	89	69	72	83
Blds-grass-drives	38	348	74	70	0	0	86
Stone-steel-towers	9	84	99	100	0	0	99
OA			71.3	68.5732	45.0547	46.7013	62.8534
AA			71	68	29	29	61
Kappa			0.6674	0.6341	0.3186	0.3531	0.5684
Parameter grid			C=0.05	C=0.02	C=0.05	n_estimators 700	Neigh=11
			G=0.01	G=0.05	G=0.01	max_features log2	

Table 4: Indian Pines Dataset with train size=30%

Class	**Train**	**test**	**SVM-l**	**SVM-p**	**SVM-r**	**RF**	**kNN**
Alfalfa	13	33	81	43	0	0	0
Corn-notill	428	1000	80	73	59	54	51
Corn-mintill	249	581	76	66	0	0	69
Corn-mintill	71	166	71	64	0	0	48
Grass-pasture	144	339	85	87	0	0	81
Grass-trees	219	511	92	87	55	35	72
Grass-pasture-mowed	8	20	100	100	0	0	0
Hay-windrowed	143	335	94	89	78	0	84
Oats	6	14	75	57	0	0	0
Soybean-notill	291	681	75	80	0	0	55
Soybean-mintill	736	1719	68	66	41	42	61
Soybean-clean	177	416	67	69	0	0	67
Wheat	61	144	95	97	91	0	79
Woods	379	886	91	93	70	72	86
Blds-grass-drives	115	271	85	71	0	0	82
Stone-steel-towers	27	66	100	100	0	0	100
OA			78.7524	76.3018	51.11	46.8671	66.7919
AA			79	76	36	29	67
Kappa			0.7544	0.7267	0.4039	0.3544	0.6146
Parameter grid			C=0.05 G=0.01	C=0.05 G=0.05	C=0.05 G=0.05	n_estimators 700 max_features log2	Neigh=10

Table 5: Indian Pines Dataset with train size=50%

Class	**Train**	**test**	**SVM-l**	**SVM-p**	**SVM-r**	**RF**	**kNN**
Alfalfa	23	23	100	50	0	0	67
Corn-notill	714	714	81	79	60	0	51
Corn-mintill	415	415	82	71	0	0	66
Corn-mintill	118	119	75	64	0	0	42
Grass-pasture	241	242	89	93	89	0	84
Grass-trees	365	365	93	93	65	35	74
Grass-pasture-mowed	14	14	100	100	0	0	67
Hay-windrowed	239	239	95	92	80	0	86
Oats	10	10	100	50	0	0	0
Soybean-notill	486	486	76	84	0	0	58
Soybean-mintill	1227	1228	75	73	41	38	63
Soybean-clean	296	297	72	78	0	0	64
Wheat	102	103	95	94	88	0	84
Woods	632	633	92	94	71	73	86
Blds-grass-drives	193	193	86	78	100	0	87
Stone-steel-towers	46	47	100	98	100	0	100
OA			82.58	81.2597	53.95	42.84	67.6321
AA			83	81	46	21	68
Kappa			0.7997	0.7845	0.4403	0.2979	0.6263
Parameter grid			C=0.05 G=0.01	C=0.05 G=0.05	C=0.05 G=0.01	n_estimators 700 max_features log2	

Table 6: Indian Pines Dataset with test size=0.4 (train size 60% of samples)

Class	**Train**	**test**	**SVM-l**	**SVM-p**	**SVM-r**	**RF**	**kNN**	**GMM**
Alfalfa	27	19	100	94	0	0	100	14
Corn-notill	856	572	83	91	59	0	61	29
Corn-mintill	498	332	87	84	97	0	73	39
Corn-mintill	142	95	74	81	0	0	55	18
Grass-pasture	289	194	92	95	91	0	88	61
Grass-trees	438	292	94	96	72	34	80	77
Grass-pasture-mowed	16	12	100	100	0	0	86	27
Hay-windrowed	286	192	96	96	82	0	87	91
Oats	12	8	100	70	0	0	80	3
Soybean-notill	583	389	80	92	65	0	70	31
Soybean-mintill	1473	982	76	75	46	38	72	58
Soybean-clean	355	238	73	88	0	0	78	15
Wheat	123	82	98	95	86	0	84	73
Woods	759	506	91	95	78	73	90	85
Blds-grass-drives	231	155	87	80	100	0	75	19
Stone-steel-towers	55	38	100	100	100	0	100	100
OA			83.8772	86.4831	60.496	30	75.45	41.65
AA			84	87	62	9	75	52
Kappa			0.8147	0.8444	0.5267	0.3	0.7178	0.3417
Parameter grid			C=0.05 G=0.01	C=0.05 G=0.05	C=0.05 G=0.02	est =700 f=log2	neigh=9	comp=9 cov=full

4 Classification Tables of Pavia Dataset

Table 7: Indian Pines Dataset with train size=80%

Class	**Train**	**test**	**SVM-l**	**SVM-p**	**SVM-r**	**RF**	**kNN**	**GMM**
Alfalfa	36	10	91	83	0	0	86	91
Corn-notill	1142	286	80	67	59	11	68	80
Corn-mintill	664	166	73	67	0	6	62	73
Corn-mintill	189	48	78	67	0	2	51	78
Grass-pasture	386	97	87	87	0	0	84	87
Grass-trees	584	146	92	84	34	0	93	92
Grass-pasture-mowed	22	6	83	56	0	0	50	83
Hay-windrowed	382	96	99	95	0	0	95	99
Oats	16	4	57	67	0	0	50	57
Soybean-notill	777	195	79	72	0	0	73	79
Soybean-mintill	1964	491	79	74	42	51	73	79
Soybean-clean	474	119	84	70	0	24	54	84
Wheat	164	41	98	83	0	0	93	98
Woods	1012	253	95	92	73	0	91	95
Blds-grass-drives	308	78	81	80	0	0	64	81
Stone-steel-towers	74	19	100	100	0	0	79	100
OA			84.62	89		46.86	74	10.96
AA			82	87.13		47	77	15
Kappa			0.81	0.861		0.3545	0.7393	0.035
Parameter grid						n_estimators 700 max_features log2		

Table 8: Indian Pines Dataset PCA (components=2)

Class	**Train**	**test**	**SVM-l**	**SVM-p**	**SVM-r**	**RF**	**kNN**
Alfalfa	27	19	0	0	0	0	57
Corn-notill	856	572	42	0	0	41	45
Corn-mintill	498	332	0	0	0	0	40
Corn-mintill	142	95	0	0	0	0	39
Grass-pasture	289	194	22	0	0	0	65
Grass-trees	438	292	64	0	47	38	71
Grass-pasture-mowed	16	12	0	0	0	0	0
Hay-windrowed	286	192	76	0	74	0	83
Oats	12	8	0	0	0	0	0
Soybean-notill	583	389	0	0	0	0	48
Soybean-mintill	1473	982	37	24	36	43	57
Soybean-clean	355	238	0	0	0	0	33
Wheat	123	82	72	0	0	0	80
Woods	759	506	71	0	65	73	80
Blds-grass-drives	231	155	0	0	0	0	38
Stone-steel-towers	55	38	0	0	0	0	40
OA			48.733	23.91	43.47	46.7121	58.5728
AA			34	6	23	28	56
Kappa			0.3716	0	0.2998	0.3533	0.5224
Parameter grid			C=0.01 G=0.01	C=0.01 G=0.01	C=0.05 G=0.05		Neighbour=12

Table 9: Indian Pines Dataset PCA (components=10)

Class	**Train**	**test**	**SVM-l**	**SVM-p**	**SVM-r**	**RF**	**kNN**
Alfalfa	27	19	18	16	0	0	9
Corn-notill	856	572	68	60	59	50	43
Corn-mintill	498	332	30	37	0	0	23
Corn-mintill	142	95	29	25	0	0	9
Grass-pasture	289	194	52	64	68	0	39
Grass-trees	438	292	87	86	68	56	70
Grass-pasture-mowed	16	12	39	2	0	0	7
Hay-windrowed	286	192	96	93	81	80	92
Oats	12	8	5	3	0	0	2
Soybean-notill	583	389	35	36	0	0	29
Soybean-mintill	1473	982	62	72	43	39	57
Soybean-clean	355	238	26	36	0	0	17
Wheat	123	82	51	77	89	0	56
Woods	759	506	87	89	67	73	82
Blds-grass-drives	231	155	58	42	100	0	15
Stone-steel-towers	55	38	75	77	97	0	99
OA			54.19	51.2439	54.55	49.97	38.24
AA			60	63	45	33	50
Kappa			0.4831	0.4572	0.4493	0.3886	0.3133
Parameter grid			C=0.05 G=0.01	C=0.02 Gamma=0.05	C=0.05 Gamma=0.05		Neighbour=8

Table 10: Indian Pines LDA (components=7)

Class	**Train**	**test**	**SVM-l**	**SVM-p**	**SVM-r**	**RF**	**kNN**
Alfalfa	27	19	0	75	64	0	54
Corn-notill	856	572	0	53	52	46	48
Corn-mintill	498	332	0	0	33	0	24
Corn-mintill	142	95	0	52	59	0	51
Grass-pasture	289	194	0	48	62	0	57
Grass-trees	438	292	41	68	69	51	70
Grass-pasture-mowed	16	12	0	0	80	0	60
Hay-windrowed	286	192	0	89	90	66	88
Oats	12	8	0	0	0	0	0
Soybean-notill	583	389	0	56	56	0	55
Soybean-mintill	1473	982	34	48	53	42	63
Soybean-clean	355	238	0	24	45	0	47
Wheat	123	82	23	82	72	0	79
Woods	759	506	71	73	71	61	78
Blds-grass-drives	231	155	0	79	53	0	27
Stone-steel-towers	55	38	0	100	100	0	100
OA			41.0375	58.2318	60.4724	48.4413	61.2274
AA			20	53	58	31	58
Kappa			0.2752	0.5038	0.534	0.3735	0.5531
Parameter grid			C=0.05 G=0.01	C=0.5 G=0.5	C=1 G=1		Neighbours=9

Table 11: Salinas Dataset with train size 2%

Class	Train	test	SVM-l	SVM-p	SVM-r	RF	kNN
Brocoli	40	1969	100	100	100	0	99
Brocoli	74	3652	98	97	98	63	97
Fallow	39	1937	92	91	90	0	86
Fallow	27	1367	98	98	98	0	97
Stubble	53	2625	95	96	96	0	92
Celery	79	3880	100	100	100	29	100
Grapes	71	3503	99	99	97	86	94
Soil_vineyard	225	11046	68	64	66	50	65
Corn green weeds	124	6079	98	96	97	86	91
Lettuce	65	3213	86	86	78	32	83
Lettuce	21	1047	91	85	96	0	92
Lettuce	38	1889	96	97	93	0	82
Lettuce	18	898	94	97	95	0	82
Lettuce	21	1049	97	95	98	0	92
Vineyard	145	7123	78	72	68	0	55
Vineyard	36	1771	99	99	99	0	99
OA			87.1939	84.99	85.29	53.35	82.0707
AA			88	86	86	34	82
Kappa			0.8564	0.8313	0.8352	0.4626	0.7995
Parameter grid			C=0.05 G=0.01	C=0.05 G=0.05	C=1 G=0.05	n_estimators max_features	

Table 12: Salinas Dataset with train size=10 samples

Class	Train	test	SVM-l	SVM-p	SVM-r	RF	kNN
Brocoli	10	1999	95	98	91	89	77
Brocoli	10	3716	99	98	99	100	99
Fallow	10	1966	83	79	77	78	68
Fallow	10	1384	97	98	98	96	96
Stubble	10	2668	96	97	97	89	88
Celery	10	3949	100	100	100	98	100
Grapes	10	3569	95	94	95	88	89
Soil_vineyard	10	11261	64	61	64	53	65
Corn green weeds	10	6193	99	98	98	92	93
Lettuce	10	3268	76	55	60	99	74
Lettuce	10	1058	74	75	64	58	56
Lettuce	10	1917	94	91	89	85	81
Lettuce	10	906	94	88	90	72	61
Lettuce	10	1060	76	43	90	37	73
Vineyard	10	7258	44	42	45	41	42
Vineyard	10	1797	98	98	94	58	98
OA			80.2182	76.7162	78.18	70.48	73.5885
AA			80	77	79	74	76
Kappa			0.7803	0.7425	0.7582	0.6783	0.709
Parameter grid			C=0.2 G=0.01	C=0.01	C=1 G=0.05	n_estimators 700 max_features log2	Neigh=8

Table 13: Salinas Dataset with train size 10%

Class	**Train**	**test**	**SVM-l**	**SVM-p**	**SVM-r**	**RF**	**kNN**
Brocoli	200	1809	100	100	100	0	100
Brocoli	372	3354	99	99	99	58	98
Fallow	197	1779	96	96	92	0	87
Fallow	139	1255	99	99	98	0	97
Stubble	267	2411	97	97	97	0	94
Celery	395	3564	100	100	100	86	100
Grapes	357	3222	100	100	99	82	96
Soil_vineyard	1127	10144	79	78	72	36	67
Corn green weeds	620	5583	99	98	99	89	95
Lettuce	327	2951	97	94	90	88	87
Lettuce	106	962	98	96	92	98	93
Lettuce	192	1735	99	97	95	0	90
Lettuce	91	825	93	95	94	0	89
Lettuce	107	963	97	94	98	0	95
Vineyard	726	6542	82	76	81	0	60
Vineyard	180	1627	100	100	99	0	99
OA			92.09	90.7051	89.1741	54.2482	84.7535
AA			92	91	90	41	84
Kappa			0.9117	0.8963	0.8788	0.4589	0.8296
Parameter grid			C=0.01	C=0.01	C=1	n_estimators 700	Neigh=9
			G=0.1	G=0.2	G=0.01	max_features auto	

Table 14: Salinas Dataset with train size 30%

Class	**Train**	**test**	**SVM-l**	**SVM-p**	**SVM-r**	**RF**	**kNN**
Brocoli	602	1407	100	100	100	0	100
Brocoli	1117	2609	100	100	99	62	99
Fallow	592	1384	99	93	93	0	91
Fallow	418	976	99	99	98	0	97
Stubble	803	1875	99	98	98	0	97
Celery	1187	2772	100	100	100	76	100
Grapes	1073	2506	100	100	100	78	98
Soil_vineyard	3381	7890	79	54	73	36	70
Corn green weeds	1860	4343	99	99	99	89	98
Lettuce	983	2295	98	96	91	90	89
Lettuce	320	748	99	96	94	100	93
Lettuce	578	1349	100	96	96	0	94
Lettuce	274	642	98	0	95	0	94
Lettuce	321	749	99	98	98	0	98
Vineyard	2180	5088	83	0	84	0	63
Vineyard	542	1265	100	99	99	0	99
OA			92.88	81.28	90.09	53.9289	86.94
AA			93	74	91	40	87
Kappa			0.9206	0.7875	0.8892	0.4554	0.8543
Parameter grid			C=1			n_estimators 700	Neigh=9
			G=0.01			max_features log2	

Table 15: Salinas Dataset with train size 50%

Class	Train	test	SVM-l	SVM-p	SVM-r	RF	kNN
Brocoli	1004	1005	100	100	100	0	100
Brocoli	1863	1863	100	100	99	66	99
Fallow	988	988	99	99	91	0	91
Fallow	697	697	99	99	97	0	97
Stubble	1339	1339	99	99	97	0	97
Celery	1979	1980	100	100	100	54	100
Grapes	1789	1790	100	100	98	78	98
Soil_vineyard	5635	5636	80	86	70	39	70
Corn green weeds	3101	3102	100	99	98	89	98
Lettuce	1639	1639	98	97	89	87	89
Lettuce	534	534	100	99	93	100	93
Lettuce	963	964	100	99	94	0	94
Lettuce	458	458	100	98	94	0	94
Lettuce	535	535	99	98	98	0	98
Vineyard	3634	3634	84	81	63	0	63
Vineyard	903	904	100	100	99	0	99
OA			93.29	93.93	86.94	54.01	86.94
AA			93	94	87	39	87
Kappa			0.9251	0.9323	0.8543	0.4598	0.8543
Parameter grid			C=1 G=0.01	C=0.2 G=0.2	C=0.2 G=0.2	n_estimators 700 max_features auto	Neigh=9

Table 16: Salinas Dataset with test size=0.4 (train size 60% of samples)

Class	Train	test	SVM-l	SVM-p	SVM-r	RF	kNN
Brocoli	1205	804	100	100	100	0	100
Brocoli	2235	1491	100	100	99	0	99
Fallow	1185	791	99	100	92	0	95
Fallow	836	558	100	99	97	0	98
Stubble	1606	1072	99	100	99	0	99
Celery	2375	1584	100	100	100	18	100
Grapes	2147	1432	100	100	99	74	99
Soil_vineyard	6762	4509	80	93	73	54	73
Corn green weeds	3721	2482	99	100	99	92	99
Lettuce	1966	1312	99	94	86	83	93
Lettuce	640	428	100	100	95	0	95
Lettuce	1156	771	100	100	96	0	97
Lettuce	549	367	100	99	96	0	96
Lettuce	642	428	100	93	99	0	99
Vineyard	4360	2908	84	61	81	0	72
Vineyard	1084	723	100	99	99	0	99
OA			93.43402	92.6084	89.5429	46.1911	89.6121
AA			94	93	90	33	90
Kappa			0.9266	0.9174	0.8831	0.3828	0.884
Parameter grid			C=0.05 G=0.01	C=0.05 G=0.05	C=0.05 G=0.05		Neigh=8

Table 17: Pavia Dataset with train size 2%

Class	**Train**	**test**	**SVM-l**	**SVM-p**	**SVM-r**	**RF**	**kNN**
Asphalt	132	6499	76	77	76	76	91
Meadows	372	18277	80	78	64	72	82
Gravel	41	2058	60	63	0	0	51
Trees	61	3003	92	92	100	100	94
Painted metal sheets	26	1319	97	98	0	0	96
Bare soil	100	4929	94	94	0	0	79
Bitumen	26	1304	4	12	0	0	63
Self Blocking Bricks	73	3609	63	63	62	43	72
Shadows	18	929	100	100	0	100	100
OA			78.85	77.84	65.97	67.8798	81.0766
AA			78	78	52	56	82
Kappa			0.7041	0.6885	0.4862	0.5355	0.7404
Parameter grid			C=0.05 G=0.01	C=0.05 G=0.05	C=0.05 G=0.05	n_estimators 700 max_features log2	Neigh=7

5 Changing the number of samples

We learn the model using randomly selected 10, 20, 30, 40 and 50 samples per class from training set. For small sample size classes i.e. alfaalfa (46), grass-pasture-mowed (28), oats (20) and stonesteel-towers (93), the number of training samples is set to 10 , in the case of Indian Pines dataset. Overall (OA) accuracy is computed using the all test samples.

Table 18: Pavia Dataset with train size=10 samples

Class	**Train**	**test**	**SVM-l**	**SVM-p**	**SVM-r**	**RF**	**kNN**
Asphalt	10	6621	49	0	98	52	94
Meadows	10	18639	81	89	88	79	89
Gravel	10	2089	30	27	23	24	27
Trees	10	3054	52	58	52	30	55
Painted metal sheets	10	1335	99	95	74	83	95
Bare soil	10	5019	38	93	30	33	32
Bitumen	10	1320	16	5	29	16	36
Self Blocking Bricks	10	3672	56	73	61	51	60
Shadows	10	937	100	100	100	86	99
OA			57.5739	33.6784	58.265	50.02	61.012
AA			63	67	72	59	74
Kappa			0.4612	0.2584	0.4998	0.3882	0.527
Parameter grid			C=0.02 G=0.01	C=0.05 G=0.05	C=0.01 G=0.05	n_estimators 700 max_features log2	Neighbours=7

Table 19: Pavia Dataset with train size 10%

Class	**Train**	**test**	**SVM-l**	**SVM-p**	**SVM-r**	**RF**	**kNN**
Asphalt	663	5968	77	77	76	74	92
Meadows	1864	16785	86	80	77	75	86
Gravel	209	1890	83	81	0	0	65
Trees	306	2758	90	95	95	96	96
Painted metal sheets	134	1211	99	99	99	0	98
Bare soil	502	4527	93	96	94	0	82
Bitumen	133	1197	0	17	0	0	74
Self Blocking Bricks	368	3314	73	77	63	44	77
Shadows	94	853	100	100	100	100	100
OA			84.09	81.27	77.31	70.56	85.75
AA			82	81	74	57	86
Kappa			0.782	0.7382	0.679	0.5813	0.8069
Parameter grid			C=0.05 G=0.01	C=0.05 G=0.05	C=0.05 G=0.05	n_estimators 700 max_features log2	Neigh=7

Table 20: Pavia Dataset with train size 30%

Class	**Train**	**test**	**SVM-l**	**SVM-p**	**SVM-r**	**RF**	**kNN**
Asphalt	1989	4642	80	84	76	74	94
Meadows	5594	13055	89	82	80	75	88
Gravel	629	1470	86	86	80	0	70
Trees	919	2145	93	96	96	96	97
Painted metal sheets	403	942	100	99	99	0	99
Bare soil	1508	3521	89	97	92	0	87
Bitumen	399	931	0	96	0	0	75
Self Blocking Bricks	1104	2578	80	81	64	43	80
Shadows	284	663	100	100	100	100	100
OA			87.3042	84.8899	79.58	70.43	87.9854
AA			85	86	79	57	88
Kappa			0.8279	0.7906	0.7143	0.5798	0.8376
Parameter grid			C=0.05 G=0.01	C=0.05 G=0.05	C=0.05 G=0.05	n_estimators 700 max_features log2	Neigh=7

Table 21: Pavia Dataset with train size 50%

Class	**Train**	**test**	**SVM-l**	**SVM-p**	**SVM-r**	**RF**	**kNN**
Asphalt	3315	3316	89	95	77	74	94
Meadows	9324	9325	93	97	81	75	89
Gravel	1049	1050	85	85	81	0	73
Trees	1532	1532	95	96	95	95	97
Painted metal sheets	672	673	100	100	100	0	99
Bare soil	2514	2515	90	93	90	0	90
Bitumen	665	665	86	89	100	0	76
Self Blocking Bricks	1841	1841	86	88	65	43	81
Shadows	473	474	100	100	100	100	100
OA			91.3842	94.73	80.37	70.4034	88.9486
AA			91	95	83	57	89
Kappa			0.8846	0.9301	0.7264	0.5792	0.8508
Parameter grid			C=1 G=0.01	C=1 G=0.2	C=0.05 G=0.05	n_estimators 700 max_features auto	Neigh=7

Table 22: Pavia Dataset with test size=0.4 (train size 60% of samples)

Class	Train	test	SVM-l	SVM-p	SVM-r	RF	kNN
Asphalt	3978	2653	78	17	76	74	94
Meadows	11189	7460	81	45	71	75	89
Gravel	1259	840	82	0	0	0	72
Trees	1838	1226	95	100	0	95	96
Painted metal sheets	807	538	99	95	98	0	99
Bare soil	3017	2012	91	100	83	0	90
Bitumen	798	532	98	0	0	0	78
Self Blocking Bricks	2209	1473	66	0	54	41	82
Shadows	568	379	100	0	100	100	100
OA			81.1663	46.172	71.23	69.5027	85.1152
AA			82	44	63	57	89
Kappa			0.738057622043	0.06	0.58411	0.56806	0.8897
Parameter grid			C=0.05 G=0.1	C=0.05 G=0.05	C=0.05 G=0.05		

6 Analysis

Table 23: Overall table Pines

Class Labels	SVM	Random Forest	k-NN	GMM	Adaboost	Autoencoder
Alfalfa	91	0	83	0		86
Corn-notill	80	59	67	11		68
Corn-mintill	73	0	67	6		62
Corn	78	0	67	2		51
Grass-pasture	87	0	87	0		84
Grass-trees	92	34	84	0		93
Grass-mowed	83	0	56	0		50
Hay-windrowed	99	0	95	0		95
Oats	57	0	67	0		50
Soybean-notill	79	0	72	0		73
Soybean-mintill	79	42	74	51		73
Soybean-clean	84	0	70	24		54
Wheat	98	0	83	0		93
Woods	95	73	92	0		91
Blgd-others	81	0	80	0		64
Stone-Steel-Towers	100	0	100	0		79
OA	84.62	46	74	10.96		74.89
AA	82	47	77	15		75
K	0.81	0.35	0.73	0.03		0.71

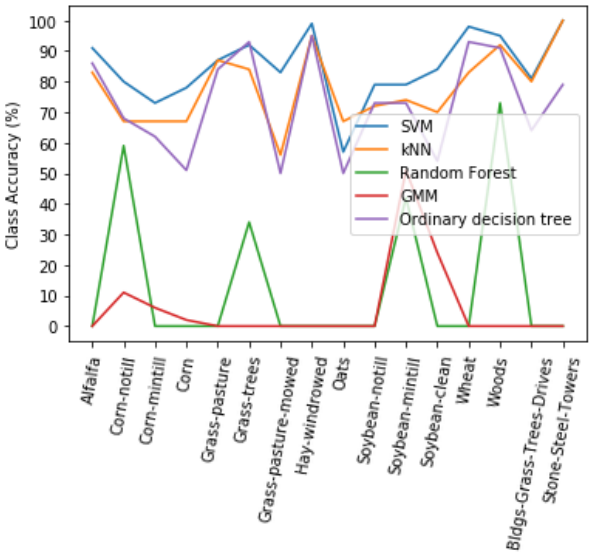
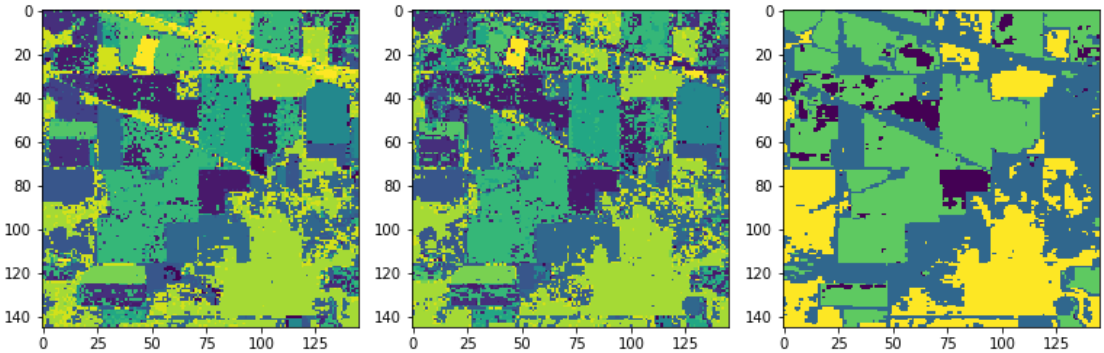
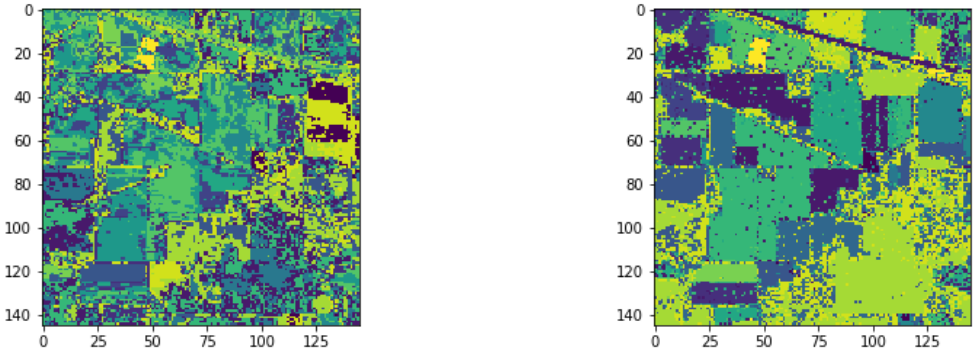


Figure 1: Classwise accuracy for different methods (Indian Pines)

Figure 2: Classification results on Indian Pines dataset



(a) SVM pines (b) kNN pines (c) Random Forest pines



(d) GMM pines (e) Decision Tree pines

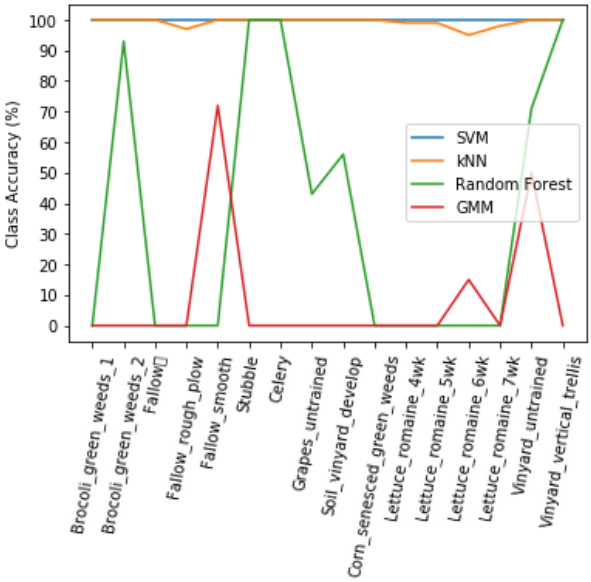
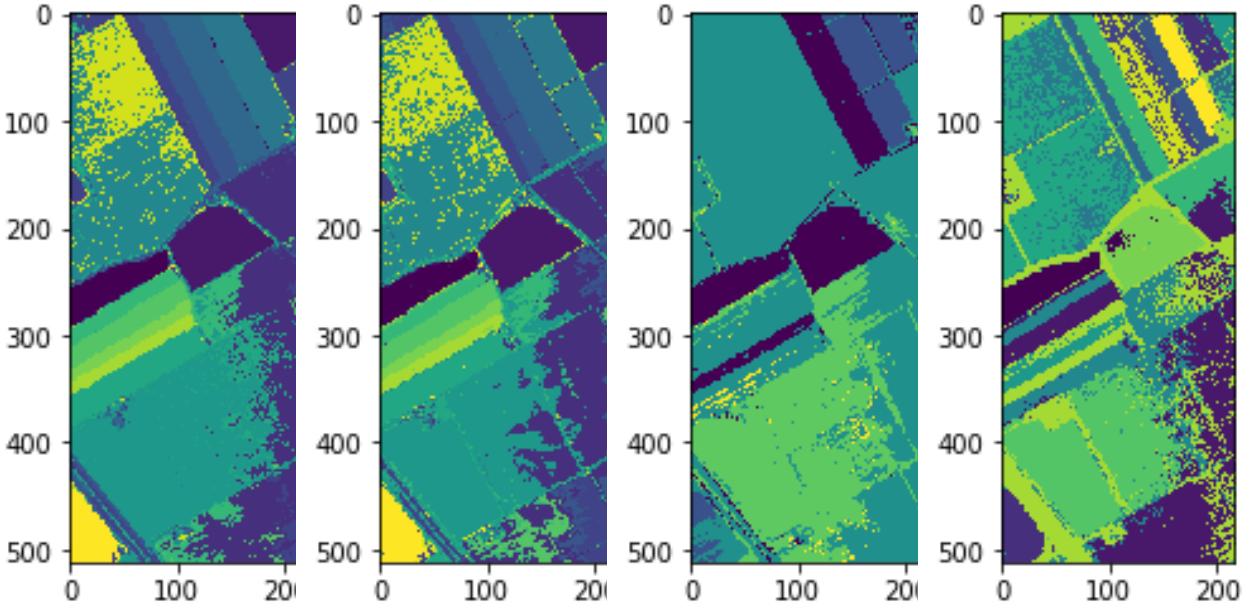


Figure 3: Classwise accuracy for different methods (Salinas)

Figure 4: Classification results on Salinas dataset



(a) SVM (b) kNN (c) Random Forest (d) GMM

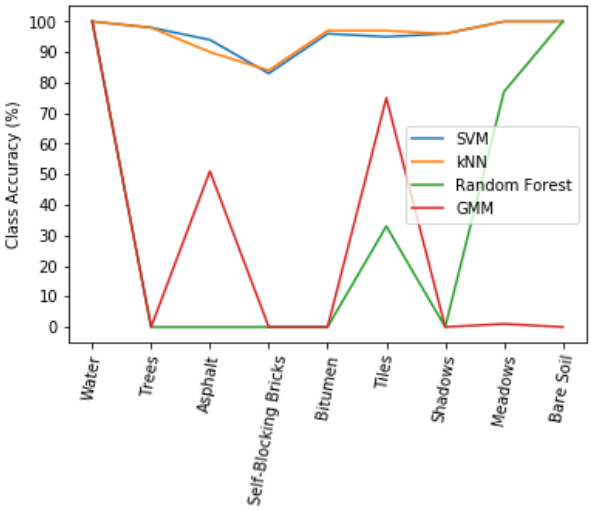
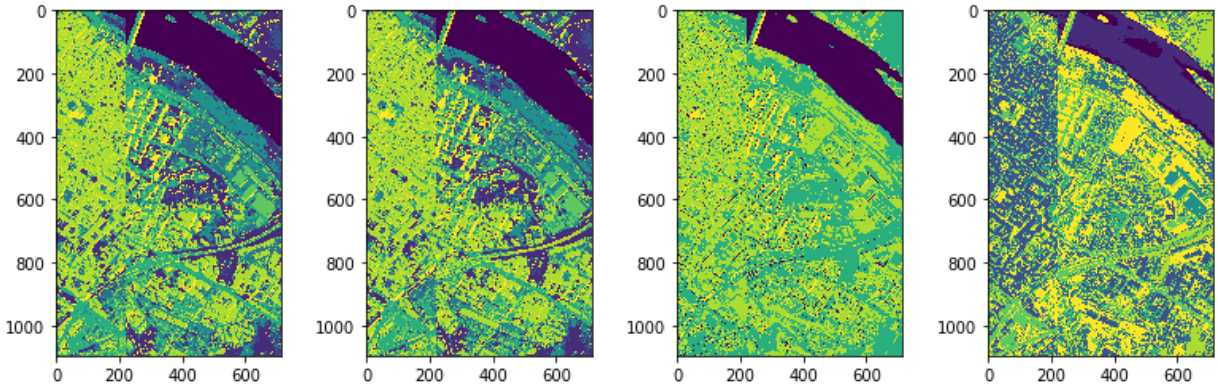


Figure 5: Classwise accuracy for different methods (Pavia)

Figure 6: Classification results on Pavia dataset



(a) SVM (b) kNN (c) Random Forest (d) GMM

Table 24: Overall table Salinas

Class Labels	SVM	Random Forest	k-NN	GMM	Adaboost
Brocoli green weeds1	100	0	100	0	
Brocoli green weeds2	100	93	100	0	
Fallow	100	0	100	0	
Fallow rough plow	100	0	97	0	
Fallow smooth	100	0	100	72	
Stubble	100	100	100	0	
Celery	100	100	100	0	
Grapes untrained	100	43	100	0	
Soil vinyard develop	100	56	100	0	
Corn greenweeds	100	0	100	0	
Lettuce romaine 4wk	100	0	99	0	
Lettuce romaine 5wk	100	0	99	0	
Lettuce romaine 6wk	100	0	95	15	
Lettuce romaine 7wk	100	0	98	0	
Vinyard untrained	100	71	100	50	
Vinyard vertical trellis	100	100	100	0	
OA	99	57.75	99	6.52	
AA	99	35.18	99.59	11	
K	0.99	0.498	0.995	.019	

Table 25: Overall table Pavia

Class Labels	SVM	Random Forest	k-NN	GMM	Adaboost
Water	100	100	100	100	
Trees	98	0	98	0	
Asphalt	94	0	90	51	
Self-Blocking Bricks	83	0	84	0	
Bitumen	96	0	97	0	
Tiles	95	33	97	75	
Shadows	96	0	96	0	
Meadows	100	77	100	1	
Bare Soil	100	100	100	0	
OA	98.71	80.06	99	46.36	
AA	99	71	98.77	51	
K	0.982	0.709	0.982	0.312	

7 Unsupervised Classification

7.1 Autoencoder

An autoencoder is an unsupervised learning algorithm that sets the target values of the neural network to be equal to the inputs. Autoencoders can get useful high-level features and can be used to learn a hierarchical feature representation. It improves in the classification by training on unlabelled data. We have used softmax classifier to classify the encoded features with the labelled samples.

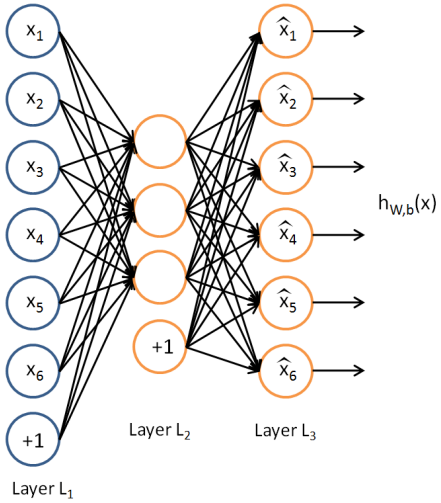


Figure 7: Image reference : http://ufldl.stanford.edu/tutorial/unsupervised/Autoencoders/

In the unsupervised classification approach, an autoencoder is trained against the Indian Pines dataset without the any class labels. The unlabelled patches are made by cutting the populated scene of Indian pines , into 8x8 patches(chosen arbitrarily). The patches are fed as input to the Autoencoder ,and then training of input image till the encoded layer is done. I used Keras library in Python to build a neural network and reduce the representation of the spectral space to a few dimensions and then upsample it to get the new features. The encoded features are fed softmax classifier which will classify the encoded features with the training set of the labelled samples. When there are large number of unlabelled samples to train on, does the model give better prediction.

Final Conclusion

All Classification methods are shown.

Code files are on the Github: `https://github.com/abunickabhi/Hyperspectral` `https://github.com/abunickabhi/autoencoder_hyperspectral` `https://github.com/abunickabhi/SVMHyperspectral`