Hamid Muneerulhudhakalvathi hxm151830@utdallas.edu

ASSIGNMENT 4 REPORT

Dataset 1:

Method	Best params	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	Avg of 10 sam.
Decision Tree	Informat--ion gain	1	1	1	1	1	1	1	1	0.955	1	0.9955
SVM	linear	0.945	0.955	0.99	0.935	0.97	0.935	0.96	0.955	0.965	0.965	0.9575
Naive Bayes	MAP classifier	0.97	0.98	0.995	0.98	0.96	0.945	0.945	0.96	0.985	0.965	0.9685
kNN	k: 7	0.845	0.81	0.85	0.87	0.83	0.84	0.82	0.865	0.835	0.835	0.84
Logistic Regression	threshold : 0.5	0.95	0.955	0.99	0.935	0.96	0.915	0.945	0.96	0.96	0.975	0.9545
Neural Net	4 hidden	0.80	0.87	0.86	0.84	0.86	0.89	0.86	0.85	0.87	0.89	0.859
Bagging	mfinal-10	1	1	1	1	1	1	1	1	0.995	1	0.9995
Random forest	default	1	1	1	1	1	1	1	1	0.995	1	0.9995
Boosting	default	1	1	1	1	1	1	1	1	0.995	1	0.9995

Dataset 2:

Method	Best params	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	Avg of 10 sam.
Decision Tree	Informat--ion gain	0.65	0.7	0.75	0.675	0.75	0.65	0.575	0.7	0.7	0.775	0.6925
SVM	linear	0.6	0.525	0.725	0.725	0.65	0.55	0.6	0.7	0.625	0.725	0.6425
Naive Bayes	MAP classifier	0.575	0.575	0.75	0.75	0.7	0.625	0.65	0.675	0.725	0.7	0.6725
kNN	k: 7	0.8	0.85	0.8	0.925	0.825	0.8	0.825	0.925	0.825	0.825	0.84
Logistic Regression	threshold : 0.5	0.6	0.575	0.75	0.7	0.7	0.625	0.6	0.675	0.65	0.725	0.66
Neural Net	4 hidden	0.6	0.525	0.725	0.725	0.7	0.55	0.6	0.7	0.675	0.725	0.6525
Bagging	mfinal-10	0.675	0.6	0.8	0.725	0.725	0.525	0.6	0.675	0.65	0.725	0.67
Random forest	default	0.65	0.675	0.75	0.75	0.75	0.625	0.6	0.7	0.65	0.75	0.69
Boosting	default	0.575	0.675	0.65	0.725	0.725	0.675	0.575	0.65	0.675	0.775	0.67

Dataset 3:

Method	Best params	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	Avg of 10 sam.
Decision Tree	Informat--ion gain	0.75	0.95	0.75	0.75	0.75	0.85	0.75	0.95	0.65	0.6	0.775
SVM	linear	0.8	0.85	0.75	0.8	0.6	0.95	0.85	1	0.8	0.75	0.815
Naive Bayes	MAP classifier	0.8	0.7	0.65	0.7	0.7	0.6	0.6	0.9	0.6	0.65	0.69
kNN	k: 7	0.8	0.85	0.9	0.75	0.75	0.9	0.65	0.8	0.75	0.7	0.785
Logistic Regression	threshold : 0.5	0.7	0.75	0.75	0.75	0.75	0.95	0.8	0.8	0.75	0.65	0.765
Neural Net	4 hidden	0.8	0.85	0.85	0.75	0.75	0.9	0.6	0.8	0.7	0.65	0.765
Bagging	mfinal-10	0.85	0.9	0.75	0.65	0.65	0.9	0.6	0.9	0.75	0.65	0.76
Random forest	default	0.85	0.85	0.8	0.8	0.75	0.9	0.65	0.9	0.65	0.65	0.78
Boosting	default	0.9	0.8	0.75	0.7	0.75	0.85	0.7	0.8	0.7	0.75	0.77

Dataset 4:

Method	Best params	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	Avg of 10 sam.
Decision Tree	Informat--ion gain	0.92	0.97	0.95	0.89	0.95	0.92	0.94	0.92	0.86	0.98	0.93
SVM	linear	0.98	1	1	0.94	0.98	0.98	0.98	0.96	0.98	0.98	0.978
Naive Bayes	MAP classifier	0.96	0.98	0.98	0.92	0.89	0.94	0.94	0.96	0.93	0.965	0.9465
kNN	k: 7	0.77	0.74	0.79	0.60	0.67	0.75	0.72	0.78	0.67	0.80	0.729
Logistic Regression	threshold : 0.5	0.98	0.97	0.98	0.94	0.96	0.92	0.96	0.92	0.98	0.95	0.956
Neural Net	4 hidden	0.65	0.71	0.60	0.55	0.58	0.65	0.63	0.65	0.64	0.64	0.63
Bagging	mfinal-10	0.96	0.98	1	0.92	0.98	0.91	0.96	0.92	0.91	1	0.954
Random forest	default	0.98	1	1	0.96	0.94	0.96	0.96	0.96	0.92	0.98	0.966
Boosting	default	0.95	0.96	0.95	0.92	0.98	0.96	0.96	0.95	0.96	0.96	0.955

Dataset 5:

Method	Best params	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	Avg of 10 sam.
Decision Tree	Informat--ion gain	0.88	0.92	0.86	0.92	0.94	0.88	0.86	0.86	0.77	0.92	0.881
SVM	linear	0.86	0.86	0.86	0.88	0.83	0.80	0.83	0.77	0.88	0.92	0.849
Naive Bayes	MAP classifier	0.86	0.86	0.94	0.94	0.94	0.88	0.86	0.88	0.88	0.92	0.896
kNN	k: 7	0.88	0.86	0.88	0.88	0.86	0.86	0.83	0.86	0.83	0.97	0.871
Logistic Regression	threshold : 0.5	0.91	0.92	0.88	0.94	0.83	0.86	0.88	0.80	0.83	0.92	0.877
Neural Net	4 hidden	1	0.98	0.92	0.98	0.98	1	0.95	0.95	0.92	0.98	0.966
Bagging	mfinal-10	0.88	0.94	0.97	0.88	0.86	0.88	0.83	0.92	0.97	0.94	0.907
Random forest	default	0.94	0.94	0.92	0.97	0.94	0.88	0.86	0.94	0.94	0.94	0.927
Boosting	default	0.92	0.88	0.92	0.97	0.94	0.88	0.92	0.94	0.97	0.92	0.926