Network Intrusion Detection and Comparative Analysis using Ensemble Machine Learning and Feature Selection

Saikat Das, Sajal Saha, Annita Tahsin Priyoti, Etee Kawna Roy, Frederick T. Sheldon, *Senior Member, IEEE,* Anwar Haque, and Sajjan Shiva, *Fellow, IEEE*

I. Introduction

This document is meant to be used alongside the article "Network Intrusion Detection and Comparative Analysis using Ensemble Machine Learning and Feature Selection". Initially in this document, we illustrate the features vs feature selection methods for all three datasets, namely NSL-KDD, UNSW-NB15 and CICIDS2017 in TABLE I, II and III, respectively, where 1 represents a feature (in a row) is selected by a corresponding FS method (in a column) and 0 represents not selected. The last column counts the total selection by the FS methods for a single feature. The features are listed in a descending order of total count value. As we have used nine FS methods, majority wins when the count is greater than 4.5 (i.e; 5). The count values for the features below 5 are discarded from the feature list and are not shown on these tables. This majority voting technique is used to select a total of 20 out of 43, 19 out of 42, and 22 out of 80 optimized features from NSL-KDD, UNSW-NB15, and CICIDS2017 datasets, respectively. Our proposed EnFS technique significantly reduces the feature set for classification model by 53.5%, 54.8%, and 72.5 % respectively for NSL-KDD, UNSW-NB15, and CICIDS2017 dataset. The rest of this document is a collection of ROC curves and tables containing results of conducted all experiments using three above mentioned datasets.

S. Das, and S. Shiva are with the Department of Computer Science, University of Memphis, Memphis, TN, 38152 USA e-mail: {sdas1, sshiva}@memphis.edu.

S. Saha, A T. Priyoti, and A. Haque are with the Department of Computer Science, Western University, London, ON N6A 5B7 Canada e-mail: {ssaha59, apriyoti, ahaque32}@uwo.ca

E K. Roy is with the Department of Electrical Engineering, Arkansas State University, Jonesboro, AR, 72401 USA email: {eteekawn.roy}@smail.astate.edu.

F.T. Sheldon is with the Department of Computer Science, University of Idaho, Moscow, ID, 83843 USA email: {sheldon}@ieee.org.

TABLE I: 20 Extracted features for NSL-KDD dataset using EnFS

Feature	ANOVA	CHI2	LASSO	LRL1	MUTINFO	PEARSON	RF	RFE	SFPR	Count
same_srv_rate	1	1	1	1	1	1	1	1	1	9
logged_in	1	1	1	1	0	1	1	1	1	8
dst_host_srv_serror_rate	1	1	1	1	0	1	1	1	1	8
dst_host_srv_count	1	1	0	1	1	1	1	1	1	8
dst_host_serror_rate	1	1	0	1	0	1	1	1	1	7
srv_serror_rate	1	1	1	1	0	1	0	1	1	7
wrong_fragment	1	1	1	1	0	1	0	1	1	7
srv_rerror_rate	1	1	1	1	0	1	0	1	1	7
diff_srv_rate	1	1	0	0	1	1	1	1	1	7
count	1	1	0	1	0	1	1	1	1	7
dst_host_same_srv_rate	1	1	0	0	1	1	1	0	1	6
flag	1	1	0	0	1	1	1	0	1	6
service	1	1	0	0	1	1	1	0	1	6
serror_rate	1	1	0	0	0	1	1	0	1	5
dst_host_diff_srv_rate	1	1	0	0	0	1	1	0	1	5
dst_host_count	1	1	0	0	0	1	1	0	1	5
srv_diff_host_rate	1	1	0	0	0	1	1	0	1	5
rerror_rate	1	1	1	0	0	1	0	0	1	5
dst_host_srv_rerror_rate	1	1	0	0	0	1	1	0	1	5
dst_host_rerror_rate	1	1	1	0	0	1	0	0	1	5

TABLE II: 19 Extracted features for UNSW-NB15 dataset using EnFS

Feature	PEARSON	MUTINFO	CHI2	ANOVA	SFPR	LRL1	LASSO	RF	RFE	Count
dttl	1	1	1	1	1	1	1	1	1	9
ct_state_ttl	1	1	1	1	1	1	1	1	1	9
proto	1	0	1	1	1	1	1	1	1	8
ct_srv_dst	1	0	1	1	1	1	1	1	1	8
dmean	1	0	1	1	1	1	0	1	1	7
swin	1	0	1	1	1	1	1	0	1	7
is_sm_ips_ports	1	0	1	1	1	1	1	0	1	7
dload	1	0	1	1	1	1	0	1	1	7
ct_srv_src	1	0	1	1	1	0	1	1	1	7
sttl	1	1	1	1	1	0	0	1	0	6
rate	1	1	1	1	1	0	0	1	0	6
sload	1	1	1	1	1	0	0	1	0	6
dwin	1	0	1	1	1	1	0	0	1	6
service	1	0	1	1	1	1	0	0	0	5
ct_dst_ltm	1	0	1	1	1	0	0	0	1	5
ackdat	1	0	1	1	1	0	0	1	0	5
sinpkt	1	0	1	1	1	0	1	0	0	5
stcpb	1	0	1	1	1	0	0	0	1	5
dur	0	1	1	0	1	0	0	1	1	5

TABLE III: 22 Extracted features for CICIDS2017 dataset using EnFS

Feature	PEARSON	MUTINFO	CHI2	ANOVA	SFPR	LRL1	LASSO	RF	RFE	Count
Packet Length Variance	1	1	1	1	1	1	1	1	1	9
Bwd Packet Length Max	1	1	1	1	1	1	1	1	0	8
Max Packet Length	1	1	1	1	1	0	1	1	1	8
Bwd Packet Length Min	1	0	1	1	1	1	1	1	1	8
Fwd IAT Std	1	0	1	1	1	1	1	0	1	7
Fwd IAT Max	1	0	1	1	1	1	1	0	1	7
PSH Flag Count	1	0	1	1	1	0	1	0	1	6
Flow IAT Max	1	0	1	1	1	1	1	0	0	6
Bwd Packet Length Std	1	0	1	1	1	0	1	1	0	6
min_seg_size_forward	1	0	0	1	1	0	1	1	1	6
Packet Length Mean	1	1	1	1	1	0	1	0	0	6
Init_Win_bytes_forward	0	1	0	0	1	0	1	1	1	5
Init_Win_bytes_backward	0	0	0	0	1	1	1	1	1	5
Fwd Packet Length Min	0	0	0	0	1	1	1	1	1	5
FIN Flag Count	0	0	1	1	1	0	1	0	1	5
ACK Flag Count	0	0	1	0	1	1	1	0	1	5
Fwd Packet Length Std	0	0	0	1	1	1	1	0	1	5
Flow Duration	0	0	1	0	1	1	1	0	1	5
Bwd IAT Mean	0	1	0	0	1	1	1	0	1	5
Active Std	1	0	0	0	1	0	1	1	1	5
Fwd IAT Min	0	1	1	1	1	0	1	0	0	5
Flow IAT Mean	0	0	0	1	1	1	1	1	0	5

TABLE IV: Details results using NSL-KDD dataset

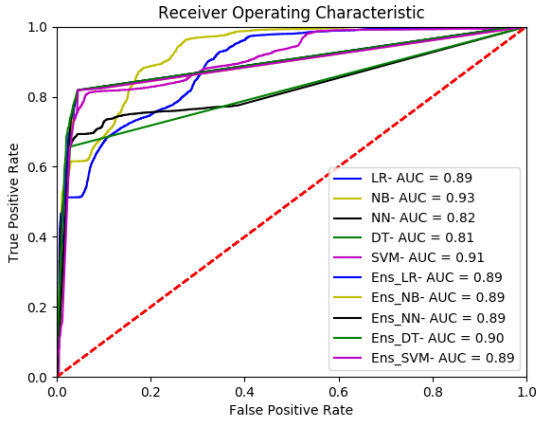
FS	Type	Classifier	F-1	Accuracy	Precision	Recall	FPR	ROC_auc	Elp_time
Full	Ind	LR	0.72	0.738	0.923	0.59	0.065	0.84	5.032
		NB	0.661	0.581	0.613	0.716	0.597	0.581	0.466
		NN	0.714	0.733	0.914	0.586	0.073	0.87	2.107
		DT	0.785	0.79	0.945	0.671	0.052	0.81	1.251
		SVM	0.73	0.754	0.971	0.585	0.023	0.874	800.549
	Ens	Ens_MV	0.73	0.745	0.92	0.605	0.069	N/A	2.352
		Ens_LR	0.813	0.811	0.936	0.718	0.065	0.881	0.294
		Ens_NB	0.81	0.805	0.91	0.73	0.095	0.803	0.201
		Ens_NN	0.818	0.817	0.944	0.721	0.056	0.843	1.355
		Ens_DT	0.825	0.823	0.945	0.732	0.057	0.889	0.209
		Ens_SVM	0.825	0.823	0.945	0.732	0.057	0.799	20.536
Anova	Ind	LR	0.732	0.746	0.915	0.611	0.075	0.889	4.272
		NB	0.751	0.768	0.965	0.615	0.03	0.927	0.31
		NN	0.786	0.795	0.967	0.662	0.029	0.824	7.654
		DT	0.781	0.791	0.968	0.655	0.029	0.814	0.754
		SVM	0.778	0.789	0.971	0.649	0.025	0.91	1029.532
	Ens	Ens_MV	0.765	0.779	0.969	0.632	0.027	N/A	2.363
		Ens_LR	0.884	0.877	0.96	0.819	0.045	0.895	0.296
		Ens_NB	0.884	0.877	0.96	0.819	0.045	0.889	0.213
		Ens_NN	0.884	0.877	0.96	0.819	0.045	0.893	0.376
		Ens_DT	0.884	0.877	0.96	0.818	0.045	0.897	0.217
		Ens_SVM	0.884	0.877	0.96	0.819	0.045	0.887	12.264
Chi2	Ind	LR	0.732	0.746	0.915	0.611	0.075	0.889	4.431
		NB	0.751	0.768	0.965	0.615	0.03	0.927	0.323
		NN	0.786	0.795	0.967	0.662	0.029	0.824	7.975
		DT	0.79	0.798	0.968	0.668	0.03	0.82	0.803
		SVM	0.778	0.789	0.971	0.649	0.025	0.91	1031.116
	Ens	Ens_MV	0.766	0.779	0.969	0.633	0.027	N/A	2.372
		Ens_LR	0.887	0.881	0.959	0.826	0.046	0.898	0.278
		Ens_NB	0.887	0.881	0.959	0.826	0.046	0.892	0.234
		Ens_NN	0.887	0.881	0.959	0.826	0.046	0.896	0.386
		Ens_DT	0.887	0.88	0.959	0.825	0.046	0.9	0.243
		Ens_SVM	0.887	0.881	0.959	0.826	0.046	0.886	12.174
Lasso	Ind	LR	0.749	0.76	0.926	0.628	0.066	0.813	3.228
		NB	0.704	0.729	0.929	0.567	0.057	0.862	0.372
		NN	0.739	0.76	0.973	0.596	0.022	0.897	14.892
		DT	0.783	0.786	0.931	0.675	0.067	0.853	0.837
		SVM	0.743	0.763	0.969	0.603	0.025	0.884	1328.611
	Ens	Ens_MV	0.731	0.749	0.939	0.598	0.052	N/A	2.423
		Ens_LR	0.834	0.832	0.95	0.744	0.052	0.865	0.307
		Ens_NB	0.762	0.769	0.922	0.65	0.073	0.853	0.272
		Ens_NN	0.836	0.829	0.923	0.764	0.085	0.84	0.318
		Ens_DT	0.85	0.846	0.959	0.763	0.043	0.871	0.238
		Ens_SVM	0.85	0.846	0.958	0.763	0.044	0.863	14.86
LRL1	Ind	LR	0.701	0.725	0.921	0.566	0.064	0.838	4.187
		NB	0.717	0.737	0.924	0.586	0.064	0.838	0.338
		NN	0.701	0.732	0.962	0.552	0.029	0.773	7.987
		DT	0.755	0.77	0.959	0.622	0.035	0.795	1.01
		SVM	0.729	0.752	0.964	0.586	0.029	0.844	1077.839
	Ens	Ens_MV	0.709	0.738	0.966	0.56	0.026	N/A	2.48
		Ens_LR	0.826	0.826	0.955	0.728	0.045	0.841	0.373
		Ens_NB	0.783	0.784	0.916	0.683	0.083	0.84	0.264
		Ens_NN	0.827	0.826	0.954	0.729	0.046	0.844	0.86

Continued

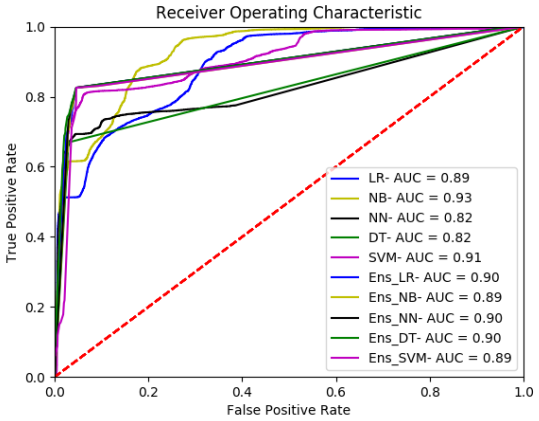
FS	Type	Classifier	F-1	Accuracy	Precision	Recall	FPR	ROC_auc	Elp_time
		Ens_DT	0.832	0.831	0.959	0.735	0.041	0.857	0.263
		Ens_SVM	0.832	0.83	0.955	0.737	0.046	0.842	19.424
MutInfo	Ind	LR	0.727	0.748	0.947	0.589	0.043	0.902	2.017
		NB	0.625	0.55	0.595	0.659	0.593	0.41	0.248
		NN	0.76	0.772	0.948	0.634	0.046	0.9	6.313
		DT	0.774	0.785	0.964	0.647	0.032	0.808	0.545
		SVM	0.773	0.782	0.95	0.651	0.046	0.82	2207.587
	Ens	Ens_MV	0.741	0.759	0.951	0.607	0.041	N/A	2.379
		Ens_LR	0.854	0.848	0.946	0.778	0.058	0.903	0.428
		Ens_NB	0.86	0.853	0.938	0.795	0.07	0.841	0.211
		Ens_NN	0.858	0.852	0.944	0.787	0.062	0.9	0.603
		Ens_DT	0.86	0.853	0.942	0.791	0.064	0.911	0.219
		Ens_SVM	0.86	0.853	0.938	0.795	0.07	0.829	16.346
Pearson	Ind	LR	0.732	0.746	0.915	0.611	0.075	0.889	4.469
		NB	0.751	0.768	0.965	0.615	0.03	0.927	0.331
		NN	0.786	0.795	0.967	0.662	0.029	0.824	7.67
		DT	0.782	0.792	0.968	0.656	0.028	0.815	0.784
		SVM	0.778	0.789	0.971	0.649	0.025	0.91	1029.37
	Ens	Ens_MV	0.765	0.779	0.969	0.632	0.027	N/A	2.39
		Ens_LR	0.884	0.877	0.96	0.818	0.045	0.895	0.3
		Ens_NB	0.884	0.877	0.96	0.818	0.045	0.889	0.245
		Ens_NN	0.884	0.877	0.96	0.818	0.045	0.893	0.362
		Ens_DT	0.883	0.877	0.96	0.818	0.045	0.896	0.25
		Ens_SVM	0.884	0.877	0.96	0.818	0.045	0.888	11.795
RF	Ind	LR	0.717	0.736	0.918	0.589	0.07	0.899	1.981
		NB	0.751	0.761	0.921	0.634	0.072	0.916	0.304
		NN	0.736	0.756	0.961	0.596	0.032	0.858	9.814
		DT	0.757	0.772	0.963	0.624	0.032	0.78	0.669
		SVM	0.706	0.736	0.966	0.556	0.026	0.838	742.154
	Ens	Ens_MV	0.73	0.753	0.967	0.587	0.026	N/A	2.385
		Ens_LR	0.816	0.811	0.918	0.734	0.087	0.843	0.302
		Ens_NB	0.789	0.788	0.913	0.694	0.087	0.839	0.259
		Ens_NN	0.818	0.819	0.953	0.717	0.047	0.843	1.455
		Ens_DT	0.817	0.818	0.954	0.715	0.046	0.846	0.237
		Ens_SVM	0.818	0.819	0.953	0.717	0.047	0.833	19.497
RFE	Ind	LR	0.69	0.718	0.92	0.552	0.064	0.83	4.702
		NB	0.729	0.745	0.922	0.603	0.067	0.828	0.326
		NN	0.718	0.744	0.962	0.572	0.03	0.808	6.563
		DT	0.722	0.747	0.963	0.577	0.029	0.774	0.846
		SVM	0.719	0.745	0.966	0.573	0.027	0.849	1094.192
	Ens	Ens_MV	0.714	0.742	0.967	0.566	0.025	N/A	2.499
		Ens_LR	0.801	0.805	0.953	0.691	0.045	0.832	0.373
		Ens_NB	0.761	0.767	0.915	0.651	0.08	0.824	0.236
		Ens_NN	0.804	0.807	0.953	0.695	0.046	0.833	1.864
		Ens_DT	0.805	0.807	0.952	0.697	0.046	0.835	0.232
		Ens_SVM	0.804	0.807	0.952	0.696	0.046	0.832	41.836
SFPR	Ind	LR	0.765	0.779	0.97	0.632	0.026	0.786	4.155
		NB	0.744	0.755	0.916	0.627	0.076	0.879	0.264
		NN	0.734	0.757	0.977	0.587	0.018	0.807	8.771
		DT	0.788	0.787	0.912	0.694	0.089	0.779	0.516
		SVM	0.735	0.757	0.968	0.592	0.026	0.865	4538.86
	Ens	Ens_MV	0.756	0.772	0.97	0.619	0.025	N/A	3.95
		Ens_LR	0.817	0.81	0.905	0.744	0.103	0.848	0.728
		Ens_NB	0.796	0.796	0.921	0.701	0.079	0.849	0.511
		Ens_NN	0.826	0.825	0.95	0.731	0.051	0.857	2.667

Continued

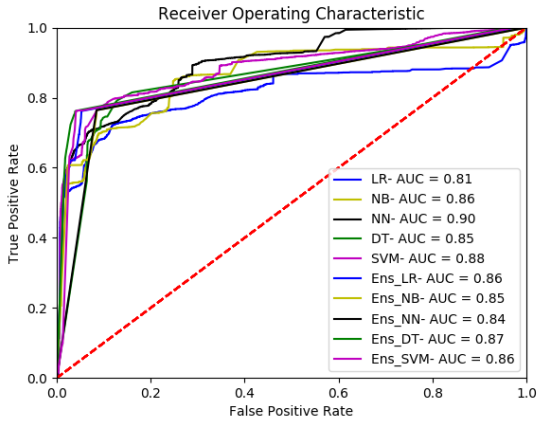
FS	Type	Classifier	F-1	Accuracy	Precision	Recall	FPR	ROC_auc	Elp_time
		Ens_DT	0.832	0.829	0.945	0.743	0.057	0.858	0.509
		Ens_SVM	0.831	0.827	0.94	0.744	0.062	0.849	48.713
EnFS	Ind	LR	0.732	0.746	0.915	0.611	0.075	0.889	4.431
		NB	0.751	0.768	0.965	0.615	0.03	0.927	0.323
		NN	0.786	0.795	0.967	0.662	0.029	0.824	7.975
		DT	0.79	0.798	0.968	0.668	0.03	0.82	0.803
		SVM	0.778	0.789	0.971	0.649	0.025	0.91	1031.116
	Ens	Ens_MV	0.766	0.779	0.969	0.633	0.027	N/A	2.372
		Ens_LR	0.887	0.881	0.959	0.826	0.046	0.898	0.278
		Ens_NB	0.887	0.881	0.959	0.826	0.046	0.892	0.234
		Ens_NN	0.887	0.881	0.959	0.826	0.046	0.896	0.386
		Ens_DT	0.887	0.88	0.959	0.825	0.046	0.9	0.243
		Ens_SVM	0.887	0.881	0.959	0.826	0.046	0.886	12.174



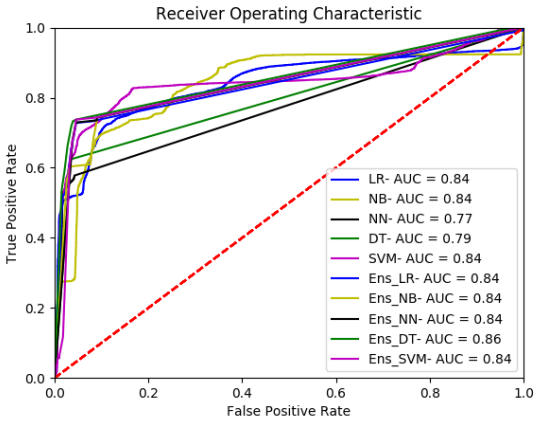
(a) Annova



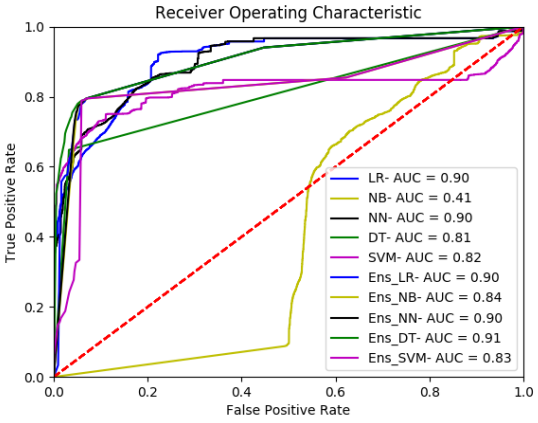
(b) Chi-Squared



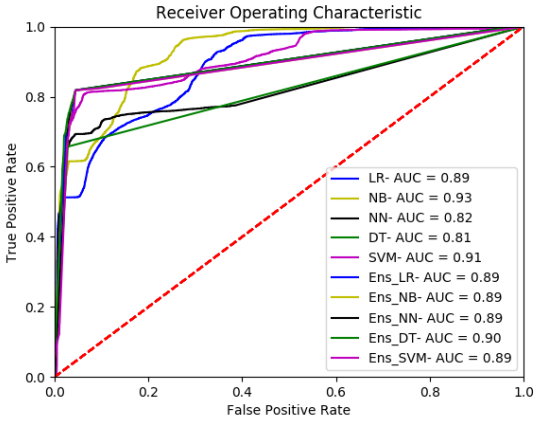
(c) Lasso



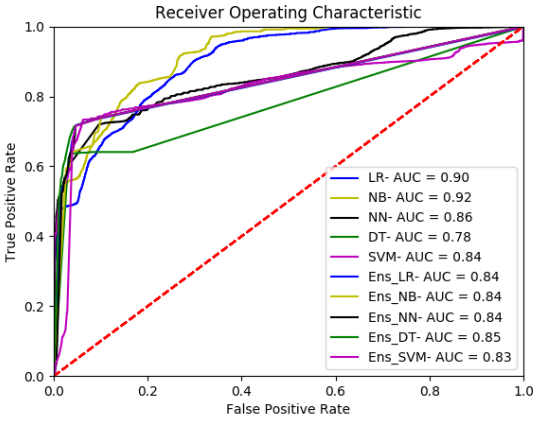
(d) LR with L1 penalty



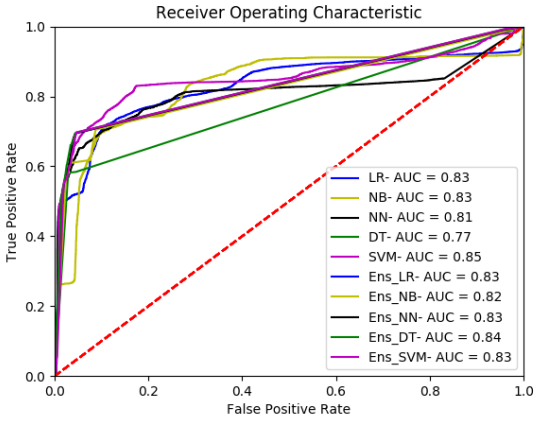
(e) Mutual Information



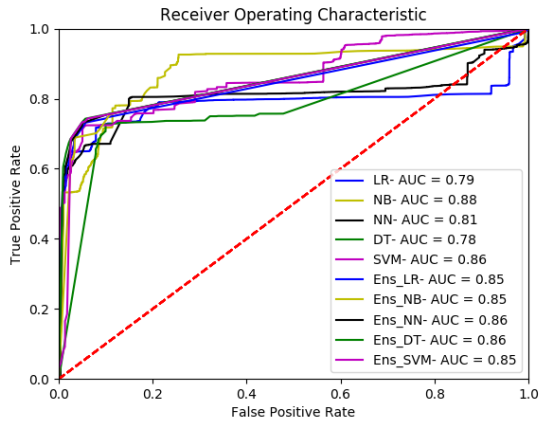
(f) Pearson



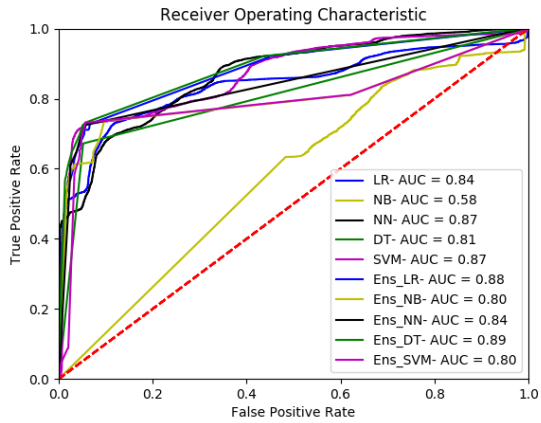
(g) Random Forest



(h) Recursive Feature Elimination



(i) Select p value using False Positive Rate



(j) Full feature set

Fig. 1: ROC Curves for nine feature selection methods and full feature set using NSL-KDD dataset.

TABLE V: Details results using UNSW-NB15 dataset

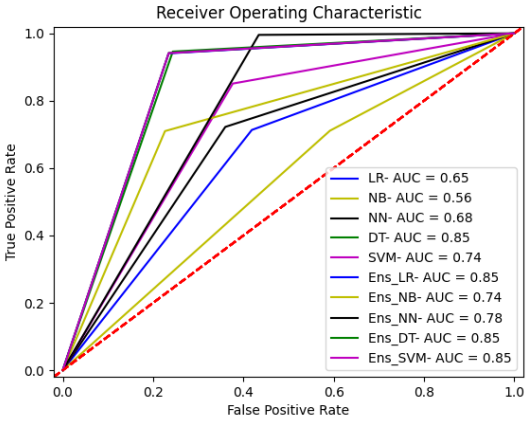
FS	Type	Classifier	F-1	Accuracy	Precision	Recall	FPR	ROC_auc	Elp_time
Full	Ind	LR	0.775	0.744	0.691	0.884	0.396	0.744	6.54
		NB	0.597	0.585	0.58	0.615	0.446	0.585	0.434
		NN	0.787	0.773	0.743	0.836	0.29	0.773	7.597
		DT	0.805	0.794	0.764	0.852	0.263	0.794	2.062
		SVM	0.826	0.796	0.719	0.97	0.379	0.796	3654.6
	Ens	Ens_MV	0.824	0.792	0.714	0.974	0.389	N/A	2.31
		Ens_LR	0.794	0.787	0.765	0.826	0.251	0.788	1.194
		Ens_NB	0.824	0.795	0.718	0.967	0.375	0.796	0.233
		Ens_NN	0.832	0.8	0.715	0.994	0.391	0.801	2.788
		Ens_DT	0.839	0.813	0.735	0.977	0.348	0.815	0.241
		Ens_SVM	0.839	0.813	0.735	0.977	0.347	0.815	273.4
Anova	Ind	LR	0.669	0.647	0.63	0.713	0.419	0.647	5.317
		NB	0.618	0.56	0.546	0.711	0.592	0.56	0.323
		NN	0.694	0.681	0.667	0.722	0.36	0.681	15.104
		DT	0.864	0.851	0.795	0.946	0.244	0.851	1.16
		SVM	0.764	0.737	0.693	0.851	0.376	0.737	3555.5
	Ens	Ens_MV	0.704	0.685	0.664	0.75	0.38	N/A	2.304
		Ens_LR	0.864	0.852	0.798	0.941	0.234	0.853	0.887
		Ens_NB	0.732	0.742	0.756	0.71	0.226	0.742	0.242
		Ens_NN	0.818	0.779	0.694	0.995	0.434	0.781	1.399
		Ens_DT	0.864	0.853	0.798	0.942	0.235	0.853	0.242
		Ens_SVM	0.864	0.853	0.798	0.941	0.235	0.853	185
Chi-2	Ind	LR	0.665	0.643	0.627	0.707	0.42	0.643	5.511
		NB	0.616	0.558	0.544	0.71	0.594	0.558	0.31
		NN	0.815	0.788	0.724	0.932	0.355	0.788	13.586
		DT	0.867	0.856	0.802	0.944	0.233	0.856	1.196
		SVM	0.762	0.735	0.692	0.847	0.376	0.735	2537.3
	Ens	Ens_MV	0.765	0.737	0.692	0.854	0.379	N/A	2.315
		Ens_LR	0.866	0.857	0.81	0.93	0.215	0.857	1.143
		Ens_NB	0.798	0.798	0.791	0.805	0.209	0.798	0.246
		Ens_NN	0.867	0.857	0.808	0.935	0.22	0.858	6.019
		Ens_DT	0.866	0.857	0.811	0.928	0.214	0.857	0.238
		Ens_SVM	0.866	0.857	0.811	0.928	0.214	0.857	174.1
Lasso	Ind	LR	0.673	0.657	0.642	0.706	0.393	0.657	5.094
		NB	0.548	0.43	0.454	0.691	0.83	0.43	0.247
		NN	0.678	0.732	0.851	0.563	0.098	0.732	12.53
		DT	0.717	0.717	0.716	0.717	0.284	0.717	0.544
		SVM	0.68	0.662	0.646	0.718	0.393	0.662	3685.4
	Ens	Ens_MV	0.677	0.66	0.644	0.714	0.394	N/A	2.298
		Ens_LR	0.803	0.787	0.743	0.874	0.297	0.788	1.044
		Ens_NB	0.793	0.778	0.739	0.854	0.297	0.778	0.255
		Ens_NN	0.664	0.497	0.497	1	1	0.5	0.336
		Ens_DT	0.805	0.789	0.743	0.879	0.3	0.79	0.253
		Ens_SVM	0.805	0.789	0.743	0.879	0.3	0.79	301.68
LRL1	Ind	LR	0.768	0.737	0.687	0.871	0.397	0.737	5.586
		NB	0.583	0.567	0.562	0.607	0.472	0.567	0.342
		NN	0.63	0.624	0.62	0.64	0.392	0.624	3.674
		DT	0.675	0.676	0.678	0.672	0.319	0.676	0.745
		SVM	0.806	0.775	0.708	0.936	0.387	0.775	4320.8
	Ens	Ens_MV	0.667	0.651	0.638	0.699	0.397	N/A	2.309
		Ens_LR	0.817	0.799	0.746	0.902	0.303	0.8	0.865
		Ens_NB	0.667	0.655	0.641	0.694	0.383	0.656	0.231
		Ens_NN	0.813	0.779	0.7	0.97	0.41	0.78	0.823

Continued

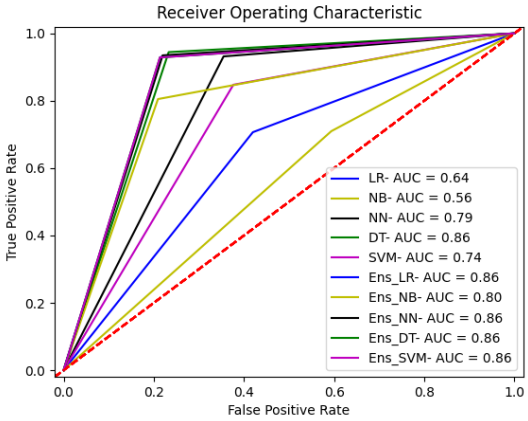
FS	Type	Classifier	F-1	Accuracy	Precision	Recall	FPR	ROC_auc	Elp_time
		Ens_DT	0.821	0.801	0.742	0.919	0.316	0.802	0.245
		Ens_SVM	0.821	0.801	0.741	0.92	0.318	0.801	254.3
MulInfo	Ind	LR	0.795	0.746	0.667	0.982	0.489	0.746	4.315
		NB	0.43	0.598	0.738	0.303	0.108	0.598	0.259
		NN	nan	0.5	nan	0	0	0.5	0.45
		DT	0.699	0.697	0.695	0.704	0.31	0.697	0.636
		SVM	0.827	0.791	0.706	0.996	0.414	0.791	4164.6
	Ens	Ens_MV	0.689	0.683	0.676	0.704	0.337	N/A	2.308
		Ens_LR	0.834	0.806	0.726	0.981	0.366	0.808	0.866
		Ens_NB	0.818	0.784	0.702	0.98	0.409	0.785	0.234
		Ens_NN	0.825	0.79	0.704	0.996	0.413	0.791	3.995
		Ens_DT	0.835	0.808	0.727	0.98	0.363	0.809	0.242
		Ens_SVM	0.835	0.808	0.727	0.98	0.363	0.809	243.3
Pearson	Ind	LR	0.669	0.647	0.63	0.713	0.419	0.647	5.319
		NB	0.618	0.56	0.546	0.711	0.592	0.56	0.311
		NN	0.694	0.681	0.667	0.722	0.36	0.681	14.957
		DT	0.863	0.85	0.794	0.946	0.246	0.85	1.193
		SVM	0.764	0.737	0.693	0.851	0.376	0.737	3667
	Ens	Ens_MV	0.704	0.685	0.663	0.751	0.381	N/A	2.301
		Ens_LR	0.863	0.851	0.797	0.94	0.236	0.852	0.925
		Ens_NB	0.731	0.741	0.753	0.711	0.229	0.741	0.247
		Ens_NN	0.828	0.795	0.711	0.99	0.397	0.797	1.344
		Ens_DT	0.863	0.852	0.797	0.942	0.237	0.852	0.246
		Ens_SVM	0.863	0.852	0.797	0.942	0.237	0.852	185.3
RF	Ind	LR	0.6	0.599	0.598	0.603	0.405	0.599	5.186
		NB	0.509	0.582	0.616	0.434	0.27	0.582	0.281
		NN	0.675	0.683	0.693	0.658	0.292	0.683	12.943
		DT	0.704	0.716	0.735	0.675	0.244	0.716	0.904
		SVM	0.672	0.662	0.651	0.695	0.372	0.662	2241.2
	Ens	Ens_MV	0.645	0.648	0.651	0.638	0.342	N/A	2.3
		Ens_LR	0.733	0.745	0.766	0.703	0.212	0.745	0.802
		Ens_NB	0.647	0.652	0.651	0.642	0.339	0.652	0.243
		Ens_NN	0.783	0.772	0.742	0.83	0.285	0.773	2.601
		Ens_DT	0.789	0.776	0.741	0.844	0.291	0.776	0.238
		Ens_SVM	0.789	0.776	0.741	0.844	0.291	0.776	315.4
RFE	Ind	LR	0.765	0.734	0.686	0.864	0.395	0.734	5.382
		NB	0.606	0.588	0.581	0.633	0.457	0.588	0.317
		NN	0.828	0.792	0.706	0.999	0.415	0.792	14.4
		DT	0.668	0.677	0.687	0.651	0.296	0.677	0.891
		SVM	0.718	0.694	0.667	0.777	0.388	0.694	4093.3
	Ens	Ens_MV	0.719	0.694	0.664	0.785	0.397	N/A	2.297
		Ens_LR	0.824	0.805	0.746	0.919	0.308	0.805	1.159
		Ens_NB	0.766	0.736	0.685	0.87	0.395	0.737	0.236
		Ens_NN	0.826	0.791	0.704	1	0.415	0.792	0.731
		Ens_DT	0.825	0.807	0.75	0.917	0.302	0.808	0.243
		Ens_SVM	0.825	0.806	0.749	0.917	0.302	0.807	236.1
SFPR	Ind	LR	0.824	0.788	0.704	0.994	0.419	0.788	5.328
		NB	0.581	0.67	0.798	0.456	0.115	0.67	0.32
		NN	0.825	0.789	0.705	0.994	0.417	0.789	6.803
		DT	0.776	0.754	0.712	0.852	0.345	0.754	1.254
		SVM	0.828	0.792	0.706	0.999	0.416	0.792	6681
	Ens	Ens_MV	0.827	0.791	0.706	0.997	0.415	N/A	2.297
		Ens_LR	0.827	0.802	0.729	0.956	0.351	0.803	1.091
		Ens_NB	0.823	0.788	0.703	0.993	0.414	0.789	0.252
		Ens_NN	0.827	0.802	0.729	0.956	0.351	0.803	2.35

Continued

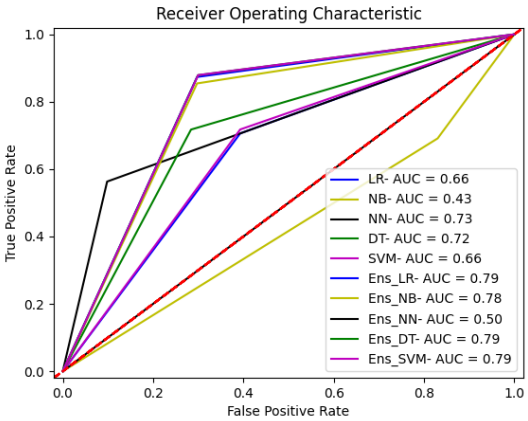
FS	Type	Classifier	F-1	Accuracy	Precision	Recall	FPR	ROC_auc	Elp_time
		Ens_DT	0.827	0.802	0.729	0.957	0.351	0.803	0.238
		Ens_SVM	0.827	0.802	0.729	0.956	0.351	0.803	250.8
EnFS	Ind	LR	0.665	0.643	0.627	0.707	0.42	0.643	5.511
		NB	0.616	0.558	0.544	0.71	0.594	0.558	0.31
		NN	0.815	0.788	0.724	0.932	0.355	0.788	13.586
		DT	0.867	0.856	0.802	0.944	0.233	0.856	1.196
		SVM	0.762	0.735	0.692	0.847	0.376	0.735	2537.3
	Ens	Ens_MV	0.765	0.737	0.692	0.854	0.379	N/A	2.315
		Ens_LR	0.866	0.857	0.81	0.93	0.215	0.857	1.143
		Ens_NB	0.798	0.798	0.791	0.805	0.209	0.798	0.246
		Ens_NN	0.867	0.857	0.808	0.935	0.22	0.858	6.019
		Ens_DT	0.866	0.857	0.811	0.928	0.214	0.857	0.238
		Ens_SVM	0.866	0.857	0.811	0.928	0.214	0.857	174.1



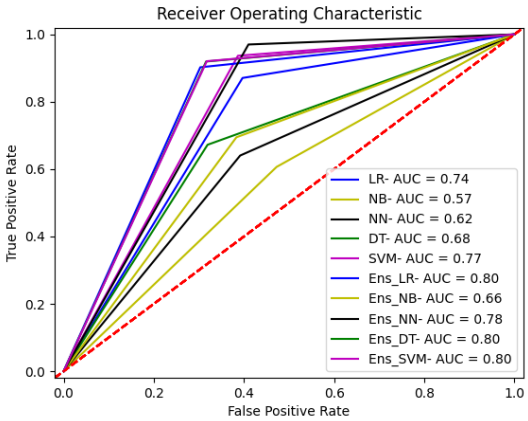
(a) Annova



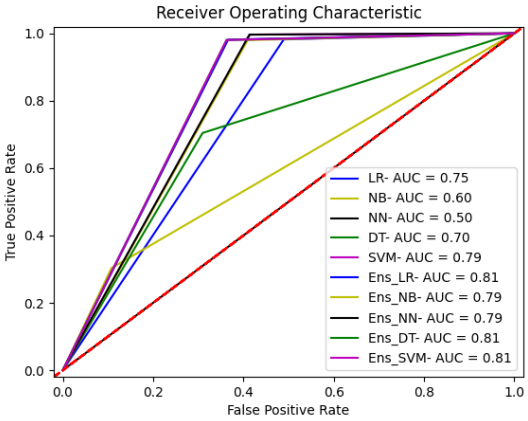
(b) Chi-Squared



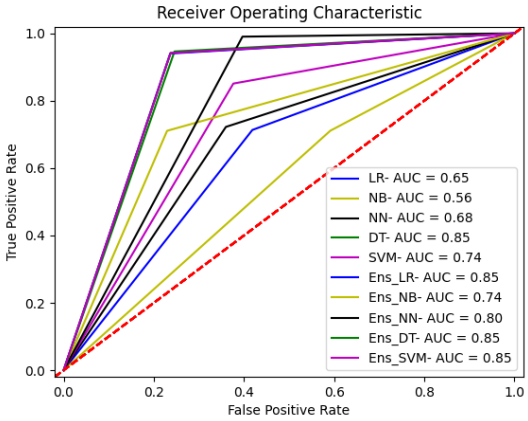
(c) Lasso



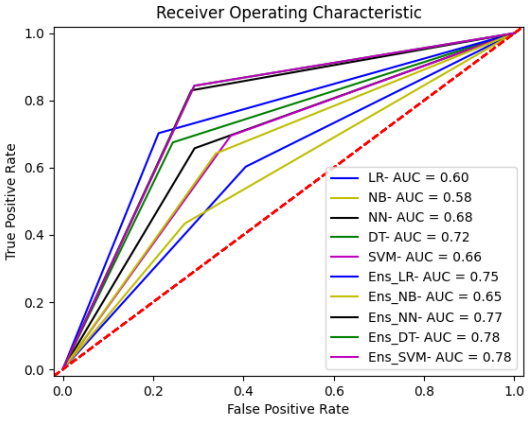
(d) LR with L1 penalty



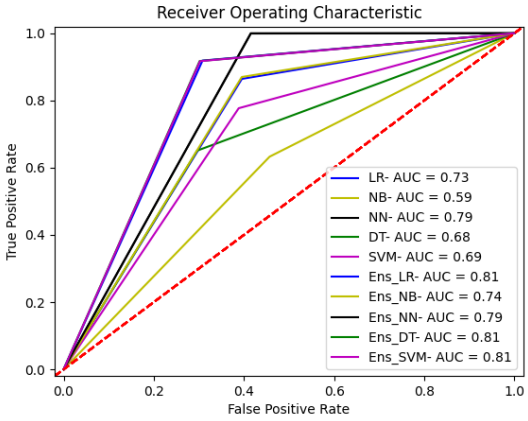
(e) Mutual Information



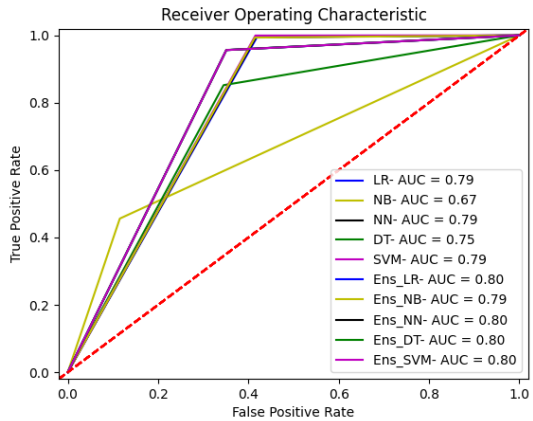
(f) Pearson



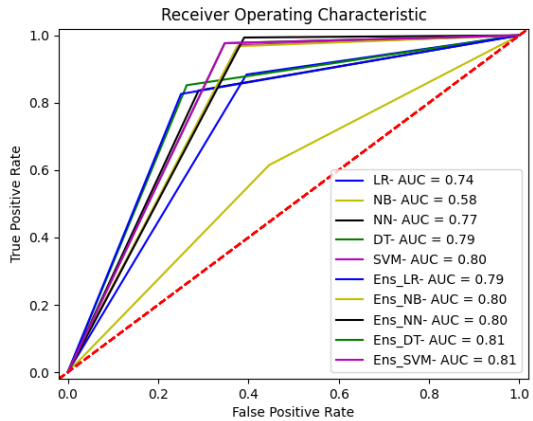
(g) Random Forest



(h) Recursive Feature Elimination



(i) Select p value using False Positive Rate



(j) Full feature set

Fig. 2: ROC Curves for nine feature selection methods and full feature set using UNSW-NB15 dataset.

TABLE VI: Details results using CICIDS2017 dataset

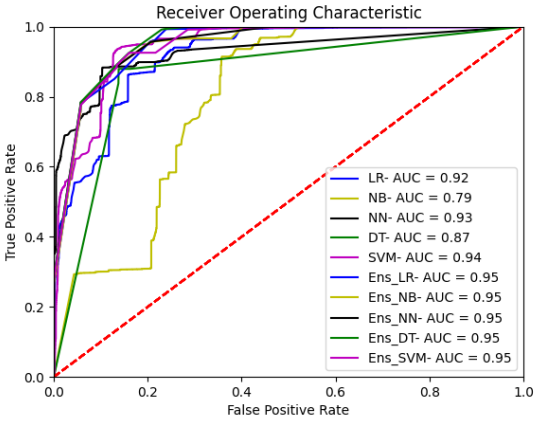
FS	Type	Classifier	F-1	Accuracy	Precision	Recall	FPR	ROC_auc	Elp_time
Full	Ind	LR	0.849	0.826	0.756	0.968	0.318	0.949	1.965
		NB	0.513	0.581	0.62	0.438	0.273	0.498	0.179
		NN	0.836	0.829	0.81	0.863	0.205	0.922	4.818
		DT	0.633	0.717	0.916	0.484	0.045	0.719	1.582
		SVM	0.893	0.88	0.811	0.994	0.236	0.928	1032.7
	Ens	Ens_MV	0.889	0.875	0.807	0.989	0.241	N/A	2.046
		Ens_LR	0.943	0.941	0.94	0.946	0.065	0.977	0.165
		Ens_NB	0.908	0.897	0.84	0.989	0.202	0.947	0.096
		Ens_NN	0.944	0.942	0.942	0.945	0.062	0.977	0.367
		Ens_DT	0.945	0.942	0.941	0.948	0.063	0.978	0.105
		Ens_SVM	0.944	0.942	0.941	0.948	0.063	0.954	3.635
Anova	Ind	LR	0.837	0.81	0.739	0.965	0.347	0.917	1.188
		NB	0.428	0.594	0.741	0.301	0.107	0.791	0.119
		NN	0.876	0.874	0.867	0.886	0.139	0.928	3.126
		DT	0.862	0.858	0.846	0.879	0.163	0.866	0.45
		SVM	0.855	0.835	0.768	0.966	0.298	0.944	395.7
	Ens	Ens_MV	0.885	0.874	0.821	0.959	0.213	N/A	2.053
		Ens_LR	0.852	0.86	0.933	0.785	0.06	0.95	0.163
		Ens_NB	0.889	0.877	0.832	0.955	0.206	0.947	0.097
		Ens_NN	0.882	0.877	0.873	0.891	0.138	0.949	0.378
		Ens_DT	0.892	0.884	0.861	0.925	0.16	0.955	0.102
		Ens_SVM	0.891	0.883	0.86	0.924	0.161	0.948	7.067
Chi-2	Ind	LR	0.855	0.832	0.757	0.983	0.322	0.917	1.139
		NB	0.435	0.595	0.735	0.309	0.113	0.768	0.116
		NN	0.866	0.853	0.802	0.941	0.236	0.936	2.888
		DT	0.933	0.929	0.892	0.979	0.121	0.971	0.446
		SVM	0.87	0.853	0.783	0.98	0.276	0.946	418.7
	Ens	Ens_MV	0.889	0.876	0.809	0.986	0.236	N/A	2.055
		Ens_LR	0.936	0.931	0.9	0.974	0.116	0.96	0.16
		Ens_NB	0.928	0.923	0.901	0.955	0.112	0.958	0.097
		Ens_NN	0.936	0.931	0.9	0.974	0.116	0.96	0.461
		Ens_DT	0.938	0.933	0.9	0.978	0.116	0.961	0.103
		Ens_SVM	0.937	0.932	0.899	0.978	0.117	0.947	4.404
Lasso	Ind	LR	0.868	0.851	0.785	0.97	0.27	0.94	1.172
		NB	0.431	0.591	0.723	0.307	0.12	0.816	0.115
		NN	0.899	0.89	0.836	0.974	0.195	0.958	2.973
		DT	0.994	0.994	0.993	0.996	0.007	0.994	0.448
		SVM	0.9	0.89	0.833	0.978	0.2	0.966	311.7
	Ens	Ens_MV	0.911	0.903	0.854	0.975	0.169	N/A	2.053
		Ens_LR	0.995	0.995	0.994	0.996	0.007	0.998	0.15
		Ens_NB	0.987	0.986	0.997	0.977	0.004	0.998	0.095
		Ens_NN	0.995	0.995	0.995	0.996	0.006	0.998	0.286
		Ens_DT	0.995	0.995	0.995	0.995	0.005	0.998	0.1
		Ens_SVM	0.995	0.995	0.994	0.996	0.007	0.993	0.537
LRL1	Ind	LR	0.841	0.821	0.76	0.941	0.302	0.886	1.162
		NB	0.557	0.641	0.739	0.446	0.161	0.801	0.112
		NN	0.858	0.841	0.781	0.951	0.271	0.905	3.183
		DT	0.936	0.933	0.899	0.977	0.112	0.974	0.645
		SVM	0.874	0.856	0.782	0.989	0.28	0.914	525.4
	Ens	Ens_MV	0.876	0.859	0.789	0.984	0.268	N/A	2.051
		Ens_LR	0.937	0.933	0.906	0.971	0.107	0.968	0.171
		Ens_NB	0.936	0.932	0.909	0.965	0.103	0.957	0.095
		Ens_NN	0.939	0.935	0.905	0.977	0.11	0.966	0.73

Continued

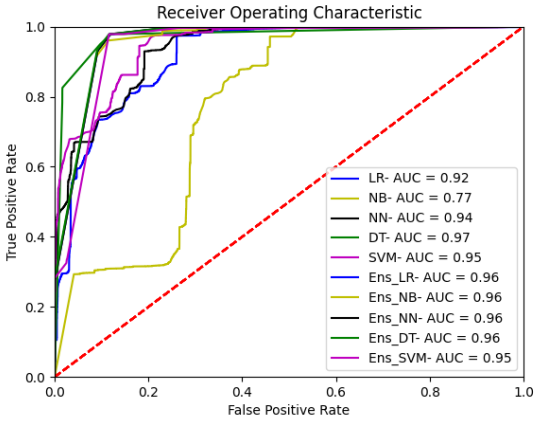
FS	Type	Classifier	F-1	Accuracy	Precision	Recall	FPR	ROC_auc	Elp_time
		Ens_DT	0.94	0.936	0.907	0.975	0.106	0.967	0.103
		Ens_SVM	0.94	0.935	0.906	0.977	0.109	0.929	4.196
MulInfo	Ind	LR	0.612	0.684	0.806	0.493	0.12	0.855	1.05
		NB	0.535	0.528	0.531	0.539	0.483	0.512	0.103
		NN	0.834	0.821	0.782	0.894	0.254	0.88	2.796
		DT	0.828	0.844	0.934	0.743	0.053	0.846	0.279
		SVM	0.855	0.835	0.767	0.966	0.299	0.904	393.8
	Ens	Ens_MV	0.832	0.817	0.777	0.894	0.261	N/A	2.055
		Ens_LR	0.943	0.943	0.959	0.928	0.042	0.967	0.145
		Ens_NB	0.869	0.851	0.798	0.953	0.258	0.928	0.097
		Ens_NN	0.945	0.944	0.964	0.926	0.037	0.964	0.282
		Ens_DT	0.945	0.944	0.965	0.925	0.036	0.97	0.101
		Ens_SVM	0.945	0.944	0.964	0.926	0.037	0.965	3.286
Pearson	Ind	LR	0.837	0.81	0.739	0.965	0.347	0.917	1.179
		NB	0.428	0.594	0.741	0.301	0.107	0.791	0.113
		NN	0.876	0.874	0.867	0.886	0.139	0.928	3.116
		DT	0.86	0.856	0.844	0.876	0.164	0.864	0.458
		SVM	0.855	0.835	0.768	0.966	0.298	0.944	380.6
	Ens	Ens_MV	0.884	0.873	0.82	0.958	0.214	N/A	2.054
		Ens_LR	0.851	0.858	0.931	0.784	0.062	0.952	0.157
		Ens_NB	0.889	0.877	0.832	0.954	0.207	0.947	0.096
		Ens_NN	0.881	0.876	0.872	0.891	0.14	0.948	0.346
		Ens_DT	0.892	0.884	0.861	0.925	0.16	0.955	0.101
		Ens_SVM	0.891	0.883	0.859	0.925	0.162	0.94	6.954
RF	Ind	LR	0.827	0.791	0.71	0.992	0.413	0.907	1.147
		NB	0.373	0.611	0.996	0.229	0.001	0.75	0.118
		NN	0.851	0.832	0.769	0.953	0.292	0.885	3.31
		DT	0.89	0.891	0.905	0.876	0.094	0.898	0.348
		SVM	0.707	0.723	0.757	0.663	0.216	0.733	384.5
	Ens	Ens_MV	0.891	0.884	0.842	0.947	0.181	N/A	2.053
		Ens_LR	0.904	0.905	0.945	0.866	0.053	0.962	0.148
		Ens_NB	0.897	0.887	0.849	0.95	0.18	0.957	0.096
		Ens_NN	0.912	0.901	0.849	0.984	0.187	0.904	0.16
		Ens_DT	0.92	0.92	0.944	0.898	0.057	0.969	0.102
		Ens_SVM	0.92	0.919	0.944	0.898	0.057	0.964	4.695
RFE	Ind	LR	0.845	0.816	0.735	0.992	0.363	0.945	1.121
		NB	0.45	0.618	0.823	0.31	0.068	0.831	0.112
		NN	0.844	0.814	0.733	0.993	0.368	0.928	2.918
		DT	0.945	0.947	0.986	0.908	0.013	0.949	0.478
		SVM	0.881	0.866	0.798	0.983	0.254	0.929	346.3
	Ens	Ens_MV	0.881	0.864	0.792	0.992	0.265	N/A	2.053
		Ens_LR	0.946	0.946	0.985	0.91	0.015	0.982	0.209
		Ens_NB	0.897	0.883	0.821	0.99	0.231	0.981	0.091
		Ens_NN	0.946	0.946	0.985	0.91	0.015	0.982	0.433
		Ens_DT	0.946	0.946	0.985	0.911	0.015	0.982	0.101
		Ens_SVM	0.946	0.946	0.985	0.91	0.015	0.973	2.817
SFPR	Ind	LR	0.843	0.814	0.734	0.991	0.366	0.86	1.128
		NB	0.427	0.59	0.722	0.303	0.119	0.774	0.115
		NN	0.867	0.852	0.792	0.958	0.256	0.927	3.02
		DT	0.931	0.928	0.897	0.968	0.113	0.97	0.547
		SVM	0.873	0.854	0.779	0.992	0.286	0.912	483.4
	Ens	Ens_MV	0.884	0.869	0.799	0.988	0.252	N/A	2.054
		Ens_LR	0.938	0.934	0.904	0.975	0.11	0.958	0.188
		Ens_NB	0.931	0.927	0.905	0.958	0.107	0.953	0.095
		Ens_NN	0.938	0.934	0.904	0.975	0.11	0.957	0.653

Continued

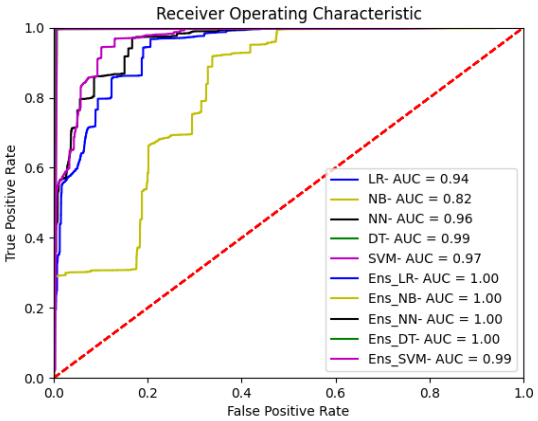
FS	Type	Classifier	F-1	Accuracy	Precision	Recall	FPR	ROC_auc	Elp_time
		Ens_DT	0.939	0.935	0.904	0.978	0.111	0.957	0.101
		Ens_SVM	0.94	0.935	0.904	0.978	0.111	0.953	4.037
EnFS	Ind	LR	0.868	0.851	0.785	0.97	0.27	0.94	1.172
		NB	0.431	0.591	0.723	0.307	0.12	0.816	0.115
		NN	0.899	0.89	0.836	0.974	0.195	0.958	2.973
		DT	0.994	0.994	0.993	0.996	0.007	0.994	0.448
		SVM	0.9	0.89	0.833	0.978	0.2	0.966	311.7
	Ens	Ens_MV	0.911	0.903	0.854	0.975	0.169	N/A	2.053
		Ens_LR	0.995	0.995	0.994	0.996	0.007	0.998	0.15
		Ens_NB	0.987	0.986	0.997	0.977	0.004	0.998	0.095
		Ens_NN	0.995	0.995	0.995	0.996	0.006	0.998	0.286
		Ens_DT	0.995	0.995	0.995	0.995	0.005	0.998	0.1
		Ens_SVM	0.995	0.995	0.994	0.996	0.007	0.993	0.537



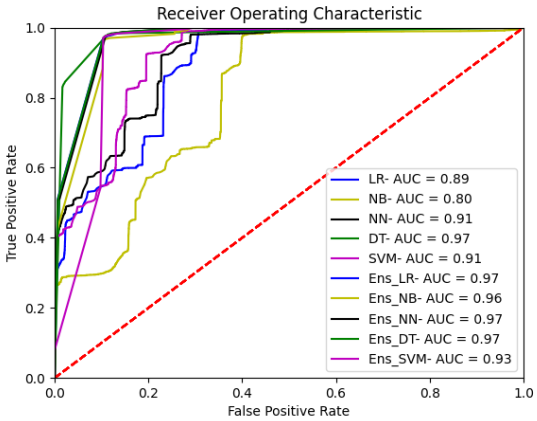
(a) Annova



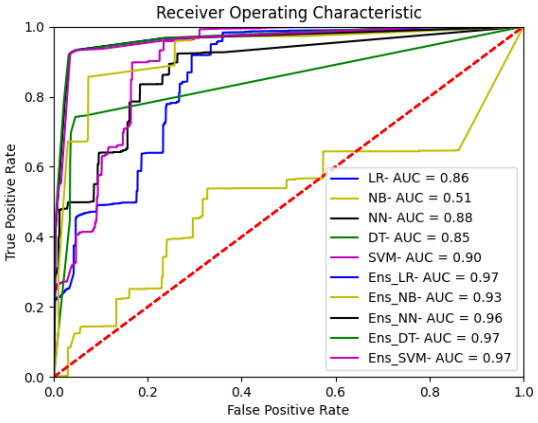
(b) Chi-Squared



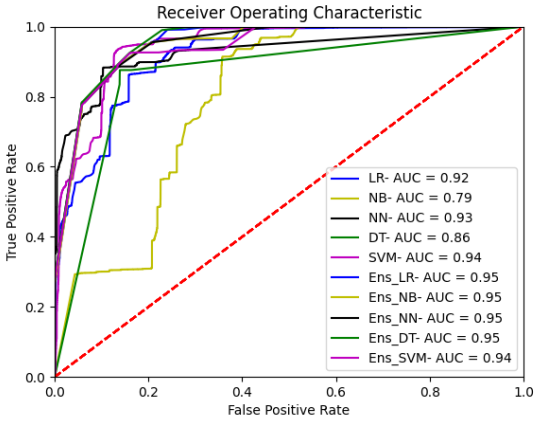
(c) Lasso



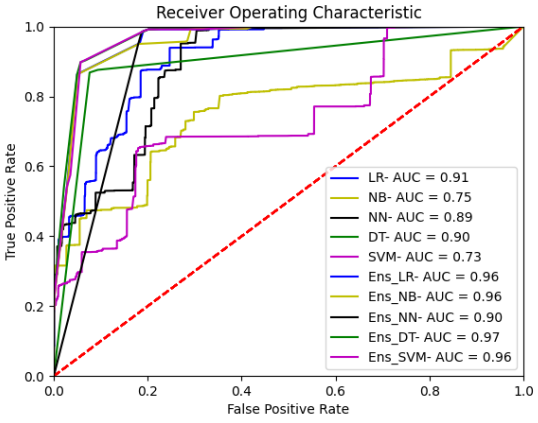
(d) LR with L1 penalty



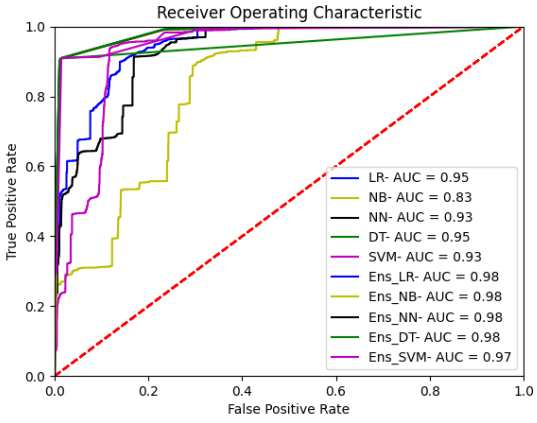
(e) Mutual Information



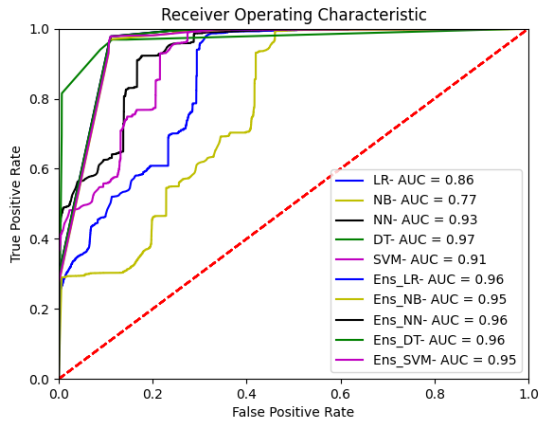
(f) Pearson



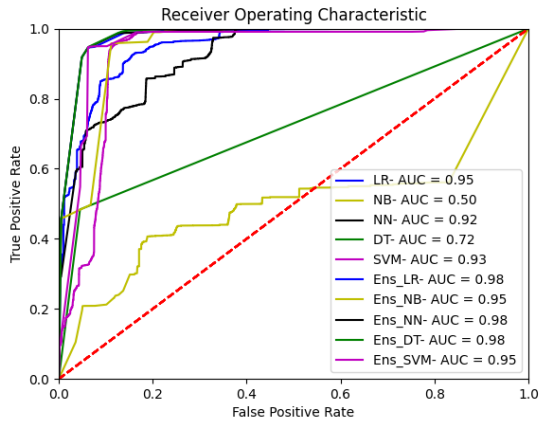
(g) Random Forest



(h) Recursive Feature Elimination



(i) Select p value using False Positive Rate



(j) Full feature set

Fig. 3: ROC Curves for nine feature selection methods and full feature set using CICIDS2017 dataset.