

1. TFIDF

1.1 intial arff

Results

=====

Correctly Classified Instances	9372	66.1771 %
Incorrectly Classified Instances	4790	33.8229 %
Kappa statistic	0.0013	
K&B Relative Info Score	191497.5127 %	
K&B Information Score	3470.1971 bits	0.245 bits/instance
Class complexity order 0	25643.9195 bits	1.8108 bits/instance
Class complexity scheme	28496.8509 bits	2.0122 bits/instance
Complexity improvement (Sf)	-2852.9314 bits	-0.2014 bits/instance
Mean absolute error	0.1062	
Root mean squared error	0.2698	
Relative absolute error	78.4711 %	
Root relative squared error	103.7253 %	
Coverage of cases (0.95 level)	84.9739 %	
Mean rel. region size (0.95 level)	42.9185 %	
Total Number of Instances	14162	

=== Detailed Accuracy By Class ===

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	
	0.000	0.000	0.000	0.000	0.000	-0.002	0.640	0.113	A
	0.000	0.000	0.000	0.000	0.000	0.000	0.627	0.049	D
	0.000	0.000	0.000	0.000	0.000	0.000	0.574	0.070	F
	0.001	0.000	0.400	0.001	0.003	0.018	0.591	0.157	H
	1.000	0.999	0.662	1.000	0.796	0.018	0.533	0.680	N
	0.003	0.000	0.667	0.003	0.005	0.039	0.564	0.113	Sa
	0.000	0.000	0.000	0.000	0.000	0.000	0.500	0.022	Su+
	0.000	0.000	0.000	0.000	0.000	0.000	0.553	0.055	Su-
Weighted Avg.	0.662	0.661	0.518	0.662	0.528	0.016	0.550	0.485	

=== Confusion Matrix ===

a	b	c	d	e	f	g	h	<-- classified as
0	0	0	0	671	0	0	0	a = A
1	0	0	0	425	0	0	0	b = D
0	0	0	1	654	0	0	0	c = F
0	0	0	2	1500	0	0	0	d = H
0	0	0	2	9368	1	0	0	e = N
0	0	0	0	787	2	0	0	f = Sa
0	0	0	0	303	0	0	0	g = Su+
0	0	0	0	445	0	0	0	h = Su-

1.2 after feature selection (cfs + greedyStepwise+startSet)

Results

=====

Correctly Classified Instances	9380	66.2336 %
Incorrectly Classified Instances	4782	33.7664 %
Kappa statistic	0.0031	
K&B Relative Info Score	31869.3412 %	
K&B Information Score	577.5161 bits	0.0408 bits/instance

Class complexity order 0	25643.9195 bits	1.8108 bits/instance
Class complexity scheme	25213.4666 bits	1.7804 bits/instance
Complexity improvement (Sf)	430.4529 bits	0.0304 bits/instance
Mean absolute error	0.134	
Root mean squared error	0.2588	
Relative absolute error	99.0587 %	
Root relative squared error	99.5038 %	
Coverage of cases (0.95 level)	97.4368 %	
Mean rel. region size (0.95 level)	84.7426 %	
Total Number of Instances	14162	

=== Detailed Accuracy By Class ===

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	
	0.004	0.000	0.750	0.004	0.009	0.056	0.549	0.119	A
	0.000	0.000	0.000	0.000	0.000	0.000	0.616	0.046	D
	0.000	0.000	0.000	0.000	0.000	0.000	0.570	0.089	F
	0.001	0.000	1.000	0.001	0.001	0.024	0.592	0.187	H
	1.000	0.998	0.662	1.000	0.797	0.037	0.470	0.649	N
	0.005	0.000	1.000	0.005	0.010	0.069	0.578	0.145	Sa
	0.000	0.000	0.000	0.000	0.000	0.000	0.518	0.028	Su+
	0.002	0.000	1.000	0.002	0.004	0.047	0.526	0.059	Su-
Weighted Avg.	0.662	0.660	0.667	0.662	0.528	0.035	0.504	0.471	

=== Confusion Matrix ===

a	b	c	d	e	f	g	h	<-- classified as
3	0	0	0	668	0	0	0	a = A
1	0	0	0	425	0	0	0	b = D
0	0	0	0	655	0	0	0	c = F
0	0	0	1	1501	0	0	0	d = H
0	0	0	0	9371	0	0	0	e = N
0	0	0	0	785	4	0	0	f = Sa
0	0	0	0	303	0	0	0	g = Su+
0	0	0	0	444	0	0	1	h = Su-

1.3 after feature selection (cfs + bestFirst)

Results

=====

Correctly Classified Instances	9375	66.1983 %
Incorrectly Classified Instances	4787	33.8017 %
Kappa statistic	0.0015	
K&B Relative Info Score	26835.5123 %	
K&B Information Score	486.2962 bits	0.0343 bits/instance
Class complexity order 0	25643.9195 bits	1.8108 bits/instance
Class complexity scheme	25291.391 bits	1.7859 bits/instance
Complexity improvement (Sf)	352.5285 bits	0.0249 bits/instance
Mean absolute error	0.1343	
Root mean squared error	0.2591	
Relative absolute error	99.2545 %	
Root relative squared error	99.6268 %	
Coverage of cases (0.95 level)	97.4862 %	
Mean rel. region size (0.95 level)	85.3905 %	
Total Number of Instances	14162	

=== Detailed Accuracy By Class ===

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	
	0.000	0.000	0.000	0.000	0.000	-0.002	0.529	0.116	A
	0.000	0.000	0.000	0.000	0.000	0.000	0.605	0.046	D
	0.000	0.000	0.000	0.000	0.000	0.000	0.561	0.083	F
	0.001	0.000	1.000	0.001	0.001	0.024	0.577	0.177	H
	1.000	0.999	0.662	1.000	0.797	0.026	0.467	0.640	N
	0.004	0.000	1.000	0.004	0.008	0.060	0.575	0.141	Sa
	0.000	0.000	0.000	0.000	0.000	0.000	0.500	0.022	Su+
	0.000	0.000	0.000	0.000	0.000	0.000	0.516	0.056	Su-
Weighted Avg.	0.662	0.661	0.600	0.662	0.528	0.023	0.499	0.463	

=== Confusion Matrix ===

a	b	c	d	e	f	g	h	<-- classified as
0	0	0	0	671	0	0	0	a = A
1	0	0	0	425	0	0	0	b = D
0	0	0	0	655	0	0	0	c = F
0	0	0	1	1501	0	0	0	d = H
0	0	0	0	9371	0	0	0	e = N
0	0	0	0	786	3	0	0	f = Sa
0	0	0	0	303	0	0	0	g = Su+
0	0	0	0	445	0	0	0	h = Su-

2. LDA

2.1 initial arff

Results

=====

Correctly Classified Instances	9371	66.17	%
Incorrectly Classified Instances	4791	33.83	%
Kappa statistic	0		
K&B Relative Info Score	183743.85	%	
K&B Information Score	3329.6901	bits	0.2351 bits/instance
Class complexity order 0	25643.9195	bits	1.8108 bits/instance
Class complexity scheme	28320.655	bits	1.9998 bits/instance
Complexity improvement (Sf)	-2676.7355	bits	-0.189 bits/instance
Mean absolute error	0.105		
Root mean squared error	0.2706		
Relative absolute error	77.5656	%	
Root relative squared error	104.062	%	
Coverage of cases (0.95 level)	86.5132	%	
Mean rel. region size (0.95 level)	46.9655	%	
Total Number of Instances	14162		

=== Detailed Accuracy By Class ===

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	
	0.000	0.000	0.000	0.000	0.000	0.000	0.580	0.062	A
	0.000	0.000	0.000	0.000	0.000	0.000	0.620	0.043	D
	0.000	0.000	0.000	0.000	0.000	0.000	0.529	0.050	F
	0.000	0.000	0.000	0.000	0.000	0.000	0.552	0.127	H
	1.000	1.000	0.662	1.000	0.796	0.000	0.518	0.674	N
	0.000	0.000	0.000	0.000	0.000	0.000	0.545	0.068	Sa
	0.000	0.000	0.000	0.000	0.000	0.000	0.523	0.023	Su+
	0.000	0.000	0.000	0.000	0.000	0.000	0.574	0.048	Su-
Weighted Avg.	0.662	0.662	0.438	0.662	0.527	0.000	0.532	0.472	

```
=== Confusion Matrix ===
```

a	b	c	d	e	f	g	h	<-- classified as
0	0	0	0	671	0	0	0	a = A
0	0	0	0	426	0	0	0	b = D
0	0	0	0	655	0	0	0	c = F
0	0	0	0	1502	0	0	0	d = H
0	0	0	0	9371	0	0	0	e = N
0	0	0	0	789	0	0	0	f = Sa
0	0	0	0	303	0	0	0	g = Su+
0	0	0	0	445	0	0	0	h = Su-

3. LSI

Numeric attribute values must all be greater or equal to zero.