

Web genre classification with methods for structured output prediction¹

Gjorgji Madjarov^a, Vedrana Vidulin^b, Ivica Dimitrovski^a, Dragi Kocev^c

^a*Faculty of Computer Science and Engineering, University Ss. Cyril and Methodius, Skopje, Macedonia*

^b*Ruđer Bošković Institute, Zagreb, Croatia*

^c*Department of Knowledge Technologies, Jožef Stefan Institute, Ljubljana, Slovenia*

Abstract

This is the supplementary information for the paper. It contains all of the results obtained with the experimental evaluation using random forests.

Keywords: web genre classification; hierarchy construction; hierarchical multi-label classification

1. Complete results using random forests

Acknowledgments

We acknowledge the financial support of the European Commission through the grant ICT-2013-612944 MAESTRA.

¹The first two authors should be regarded as joint first authors.

Email addresses: Gjorgji.Madjarov@finki.ukim.mk (Gjorgji Madjarov), Vedrana.Vidulin@irb.hr (Vedrana Vidulin), Ivica.Dimitrovski@finki.ukim.mk (Ivica Dimitrovski), Dragi.Kocev@ijs.si (Dragi Kocev)

Table 1: The performance of the different machine learning tasks applied on the different features using random forests. The hierarchy construction method are abbreviated as follows: balanced k -means clustering (BkM), predictive clustering tree (PCT), clustering with complete linkage (CL), clustering with single linkage (SL), random (RND) and manual (MAN). The evaluation measure is Hamming loss where smaller value means better result.

20-genres	HMC										HSC			MLC		SC		
	CL	SL	B2M	B3M	B4M	PCT	RND	MAN	CL	SL	B2M	B3M	B4M	PCT	RND	MAN		
BOWSrf	0.066	0.067	0.064	0.064	0.064	0.064	0.067	0.062	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.062	0.064
TFIDFSrf	0.066	0.066	0.06	0.061	0.06	0.061	0.065	0.061	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.059	0.058
Struc	0.084	0.083	0.078	0.079	0.079	0.08	0.082	0.078	0.001	0	0.001	0	0.001	0.001	0.001	0	0.078	0.091
Pres	0.078	0.079	0.078	0.078	0.078	0.079	0.079	0.077	0.001	0.001	0.001	0	0.001	0.001	0.001	0.001	0.078	0.078
Context	0.1	0.1	0.101	0.102	0.101	0.102	0.101	0.102	0.045	0.045	0.045	0.045	0.045	0.045	0.045	0.044	0.102	0.078
PV_R	0.101	0.102	0.067	0.068	0.092	0.073	0.067	0	0	0	0	0	0	0	0	0	0.065	0.089
PV_C	0.077	0.076	0.066	0.067	0.066	0.068	0.073	0.067	0	0	0	0	0	0	0	0	0.065	0.071
PV_Gl_R	0.104	0.104	0.097	0.097	0.095	0.094	0.103	0.095	0	0	0	0	0	0	0	0	0.094	0.094
PV_Gl_C	0.08	0.078	0.073	0.072	0.074	0.072	0.078	0.072	0	0	0	0	0	0	0	0	0.075	0.075
Ngrams	0.074	0.073	0.067	0.067	0.067	0.072	0.068	0.001	0.001	0	0	0	0	0	0	0	0.067	0.066
SPIRIT																		
BOWSrf	0.089	0.089	0.086	0.083	0.083	0.089	0.083	0.007	0.006	0.008	0.006	0.006	0.007	0.007	0.007	0.087	0.086	
TFIDFSrf	0.09	0.088	0.082	0.083	0.082	0.081	0.091	0.082	0.006	0.006	0.007	0.006	0.006	0.006	0.006	0.007	0.082	0.086
Struc	0.086	0.089	0.085	0.086	0.087	0.087	0.088	0.009	0.008	0.009	0.008	0.009	0.009	0.009	0.009	0.083	0.086	
Pres	0.094	0.094	0.096	0.095	0.096	0.095	0.095	0.095	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.011	0.097	0.099
Context	0.11	0.111	0.112	0.111	0.112	0.112	0.111	0.111	0.08	0.079	0.08	0.079	0.08	0.079	0.081	0.08	0.113	0.111
PV_R	0.098	0.096	0.09	0.09	0.097	0.087	0.088	0.091	0.005	0.005	0.005	0.005	0.005	0.004	0.005	0.005	0.092	0.093
PV_C	0.095	0.093	0.095	0.091	0.092	0.093	0.094	0.094	0.004	0.004	0.005	0.005	0.004	0.004	0.005	0.005	0.095	0.092
PV_Gl_R	0.103	0.106	0.105	0.105	0.105	0.104	0.107	0.004	0.004	0.004	0.004	0.005	0.004	0.004	0.005	0.004	0.108	0.098
PV_Gl_C	0.096	0.097	0.095	0.096	0.098	0.098	0.096	0.004	0.005	0.005	0.004	0.004	0.005	0.005	0.004	0.004	0.096	0.095
Ngrams	0.093	0.093	0.095	0.093	0.094	0.094	0.096	0.005	0.005	0.005	0.005	0.005	0.005	0.004	0.005	0.006	0.094	0.089

Table 2: The performance of the different machine learning tasks applied on the different features using random forests. The hierarchy construction method are abbreviated as follows: balanced k -means clustering (BkM), predictive clustering tree (PCT), clustering with complete linkage (CL), clustering with single linkage (SL), random (RND) and manual (MAN). The evaluation measure is accuracy where larger value means better result.

20-genres	HMC						HSC						MLC	SC		
	CL	SL	B2M	B3M	B4M	PCT	RND	MAN	CL	SL	B2M	B3M	B4M	PCT	RND	MAN
BOWSrf	0.375	0.362	0.444	0.444	0.441	0.448	0.374	0.455	0.994	0.993	0.996	0.993	0.993	0.995	0.476	0.414
TFIDFSrf	0.386	0.387	0.483	0.475	0.478	0.473	0.409	0.473	0.994	0.993	0.993	0.991	0.992	0.992	0.508	0.416
Struc	0.243	0.254	0.297	0.288	0.287	0.29	0.258	0.298	0.994	0.995	0.994	0.995	0.995	0.996	0.299	0.23
Pres	0.285	0.282	0.302	0.298	0.298	0.303	0.297	0.312	0.994	0.993	0.995	0.995	0.994	0.994	0.289	0.244
Context	0.249	0.247	0.273	0.26	0.265	0.269	0.251	0.263	0.533	0.532	0.535	0.528	0.536	0.53	0.531	0.265
PV_R	0.266	0.263	0.296	0.281	0.282	0.331	0.246	0.304	1	0.999	0.999	1	1	0.999	0.999	0.301
PV_C	0.27	0.275	0.371	0.366	0.373	0.36	0.299	0.368	0.999	0.999	0.998	1	0.998	0.999	0.999	0.377
PV_GI_R	0.234	0.224	0.276	0.27	0.271	0.282	0.239	0.275	1	0.99	0.99	0.999	1	1	1	0.288
PV_GI_C	0.229	0.229	0.255	0.255	0.245	0.259	0.237	0.263	0.999	1	0.999	0.999	1	1	1	0.252
Ngrams	0.31	0.318	0.402	0.407	0.402	0.407	0.336	0.399	0.993	0.994	0.996	0.996	0.996	0.995	0.995	0.372
SPIRIT																
BOWsrf	0.537	0.541	0.534	0.539	0.543	0.54	0.55	0.547	0.951	0.957	0.947	0.954	0.951	0.952	0.948	0.554
TFIDFSrf	0.544	0.546	0.543	0.542	0.545	0.551	0.538	0.548	0.958	0.963	0.951	0.961	0.959	0.956	0.954	0.546
Struc	0.539	0.526	0.543	0.54	0.534	0.539	0.535	0.531	0.936	0.941	0.938	0.939	0.945	0.94	0.933	0.55
Pres	0.495	0.497	0.49	0.494	0.49	0.488	0.496	0.502	0.926	0.926	0.925	0.919	0.924	0.92	0.923	0.52
Context	0.436	0.429	0.427	0.429	0.425	0.427	0.428	0.431	0.574	0.577	0.571	0.578	0.572	0.566	0.574	0.478
PV_R	0.508	0.514	0.508	0.511	0.514	0.514	0.513	0.504	0.963	0.96	0.959	0.962	0.961	0.969	0.959	0.423
PV_C	0.495	0.5	0.496	0.517	0.508	0.507	0.498	0.504	0.969	0.971	0.964	0.966	0.971	0.969	0.966	0.5
PV_GI_R	0.481	0.468	0.471	0.474	0.477	0.476	0.48	0.47	0.971	0.971	0.97	0.966	0.967	0.965	0.967	0.497
PV_GI_C	0.482	0.48	0.492	0.49	0.487	0.481	0.486	0.488	0.969	0.968	0.968	0.972	0.968	0.964	0.966	0.483
Ngrams	0.506	0.5	0.502	0.509	0.503	0.505	0.492	0.511	0.966	0.966	0.963	0.967	0.964	0.969	0.957	0.501

Table 3: The performance of the different machine learning tasks applied on the different features using random forests. The hierarchy construction method are abbreviated as follows: balanced k -means clustering (BkM), predictive clustering tree (PCT), clustering with complete linkage (CL), clustering with single linkage (SL), random (RND) and manual (MAN). The evaluation measure is *precision* where larger value means better result.

20-genres	HMC										HSC				MLC		SC	
	CL	SL	B2M	B3M	B4M	PCT	RND	MAN	CL	SL	B2M	B3M	B4M	PCT	RND	MAN		
BOWSrf	0.427	0.413	0.497	0.496	0.492	0.499	0.428	0.509	0.996	0.994	0.996	0.996	0.995	0.996	0.995	0.995	0.526	0.466
TFIDFSrf	0.44	0.439	0.536	0.527	0.531	0.526	0.464	0.525	0.995	0.995	0.996	0.995	0.994	0.994	0.995	0.995	0.559	0.476
Struc	0.289	0.297	0.345	0.336	0.333	0.334	0.302	0.347	0.995	0.996	0.994	0.996	0.995	0.995	0.995	0.995	0.345	0.268
Pres	0.328	0.328	0.344	0.34	0.343	0.349	0.34	0.357	0.994	0.993	0.995	0.996	0.995	0.995	0.993	0.994	0.336	0.281
Context	0.286	0.31	0.297	0.303	0.307	0.289	0.3	0.579	0.579	0.581	0.583	0.573	0.578	0.581	0.58	0.299	0.282	
PV_R	0.299	0.3	0.343	0.326	0.329	0.369	0.293	0.35	1	0.999	0.999	1	1	0.999	0.999	0.999	0.345	0.314
PV_C	0.315	0.323	0.422	0.418	0.42	0.408	0.347	0.42	1	0.999	0.999	1	0.999	0.999	1	0.999	0.424	0.365
PV_GlR	0.27	0.259	0.31	0.305	0.305	0.318	0.278	0.311	1	0.999	0.999	0.999	1	1	0.999	1	0.322	0.267
PV_GlC	0.273	0.273	0.298	0.299	0.288	0.298	0.281	0.306	0.999	1	0.999	1	1	0.998	1	1	0.294	0.296
Ngrams	0.355	0.364	0.455	0.456	0.451	0.455	0.382	0.445	0.995	0.996	0.998	0.998	0.998	0.998	0.997	0.997	0.472	0.424
SPIRIT																		
BOWSrf	0.635	0.64	0.652	0.659	0.657	0.662	0.649	0.663	0.961	0.967	0.961	0.965	0.965	0.964	0.962	0.96	0.65	0.637
TFIDFSrf	0.643	0.644	0.661	0.667	0.661	0.67	0.634	0.667	0.97	0.973	0.964	0.972	0.971	0.97	0.968	0.967	0.672	0.635
Struc	0.652	0.639	0.655	0.658	0.644	0.65	0.648	0.647	0.951	0.956	0.954	0.954	0.959	0.955	0.956	0.951	0.668	0.632
Pres	0.617	0.617	0.608	0.612	0.604	0.607	0.615	0.609	0.946	0.944	0.942	0.937	0.944	0.944	0.948	0.94	0.605	0.593
Context	0.547	0.537	0.535	0.54	0.532	0.538	0.537	0.538	0.655	0.657	0.655	0.659	0.657	0.66	0.652	0.655	0.524	0.532
PV_R	0.603	0.614	0.626	0.635	0.608	0.634	0.641	0.627	0.978	0.974	0.97	0.974	0.973	0.98	0.972	0.972	0.626	0.602
PV_C	0.608	0.617	0.611	0.625	0.617	0.609	0.612	0.979	0.983	0.976	0.977	0.981	0.978	0.975	0.978	0.615	0.6	
PV_GlR	0.592	0.576	0.573	0.589	0.587	0.589	0.587	0.581	0.982	0.983	0.98	0.98	0.98	0.98	0.977	0.976	0.983	0.574
PV_GlC	0.611	0.608	0.615	0.618	0.615	0.604	0.609	0.979	0.978	0.981	0.979	0.982	0.979	0.977	0.978	0.617	0.604	
Ngrams	0.616	0.613	0.612	0.618	0.613	0.618	0.604	0.616	0.972	0.973	0.971	0.974	0.972	0.975	0.972	0.968	0.614	0.613

Table 4: The performance of the different machine learning tasks applied on the different features using random forests. The hierarchy construction method are abbreviated as follows: balanced k -means clustering (BkM), predictive clustering tree (PCT), clustering with complete linkage (CL), clustering with single linkage (SL), random (RND) and manual (MAN). The evaluation measure is *recall* where larger value means better result.

20-genres	HMC						HSC						MLC	SC		
	CL	SL	B2M	B3M	B4M	PCT	RND	MAN	CL	SL	B2M	B3M	B4M	PCT	RND	MAN
BOWSrf	0.459	0.444	0.56	0.562	0.549	0.558	0.465	0.566	0.995	0.994	0.995	0.994	0.999	0.999	0.602	0.54
TFIDFSrf	0.478	0.476	0.601	0.588	0.591	0.595	0.503	0.591	0.995	0.995	0.996	0.994	0.993	0.994	0.642	0.518
Struc	0.327	0.328	0.389	0.376	0.373	0.375	0.342	0.342	0.999	0.999	0.998	0.999	0.999	0.999	0.378	0.35
Pres	0.344	0.341	0.368	0.364	0.368	0.366	0.359	0.359	0.999	0.998	0.999	0.999	1	0.999	0.999	0.345
Context	0.361	0.359	0.4	0.381	0.386	0.388	0.371	0.384	0.609	0.606	0.608	0.609	0.603	0.605	0.604	0.38
PV_R	0.429	0.426	0.34	0.327	0.33	0.538	0.285	0.356	1	1	1	1	1	1	0.339	0.464
PV_C	0.337	0.339	0.441	0.44	0.444	0.436	0.367	0.446	0.999	1	0.998	1	0.999	1	0.999	0.441
PV_GI_R	0.377	0.364	0.438	0.423	0.424	0.436	0.382	0.435	1	1	1	1	1	1	1	0.453
PV_GI_C	0.292	0.288	0.307	0.304	0.297	0.31	0.296	0.315	1	1	1	1	1	1	1	0.392
Ngrams	0.383	0.389	0.494	0.504	0.49	0.497	0.412	0.489	0.994	0.995	0.996	0.997	0.996	0.996	0.995	0.348
<hr/>																
SPiRiT																
BOWSrf	0.657	0.653	0.616	0.611	0.621	0.617	0.674	0.626	0.974	0.978	0.969	0.975	0.974	0.973	0.963	0.691
TFIDFSrf	0.662	0.67	0.615	0.624	0.628	0.634	0.664	0.627	0.976	0.981	0.975	0.98	0.977	0.973	0.972	0.63
Struc	0.618	0.61	0.63	0.62	0.615	0.625	0.619	0.629	0.959	0.965	0.966	0.965	0.969	0.961	0.956	0.631
Pres	0.564	0.581	0.574	0.582	0.578	0.572	0.586	0.586	0.957	0.957	0.955	0.948	0.954	0.959	0.953	0.577
Context	0.552	0.553	0.554	0.553	0.547	0.547	0.557	0.562	0.706	0.712	0.698	0.706	0.703	0.703	0.705	0.558
PV_R	0.643	0.655	0.574	0.583	0.648	0.58	0.587	0.577	0.974	0.971	0.972	0.971	0.978	0.971	0.97	0.615
PV_C	0.584	0.586	0.594	0.609	0.601	0.596	0.602	0.594	0.98	0.981	0.978	0.98	0.981	0.982	0.978	0.596
PV_GI_R	0.589	0.574	0.593	0.595	0.593	0.591	0.608	0.585	0.977	0.976	0.976	0.973	0.971	0.973	0.971	0.574
PV_GI_C	0.557	0.559	0.574	0.572	0.569	0.559	0.573	0.567	0.982	0.981	0.978	0.979	0.983	0.975	0.975	0.572
Ngrams	0.592	0.585	0.6	0.606	0.6	0.603	0.583	0.603	0.986	0.984	0.986	0.984	0.983	0.985	0.978	0.588

Table 5: The performance of the different machine learning tasks applied on the different features using random forests. The hierarchy construction method are abbreviated as follows: balanced k -means clustering (BkM), predictive clustering tree (PCT), clustering with complete linkage (CL), clustering with single linkage (SL), random (RND) and manual (MAN). The evaluation measure is F_1 score where larger value means better result.

20-genres	HMC						HSC						MLC		SC			
	CL	SL	B2M	B3M	B4M	PCT	RND	MAN	CL	SL	B2M	B3M	B4M	PCT	RND	MAN		
BOWSrf	0.442	0.428	0.527	0.527	0.519	0.527	0.446	0.536	0.996	0.994	0.995	0.997	0.995	0.995	0.997	0.997	0.561	0.5
TFIDFSrf	0.458	0.457	0.567	0.556	0.56	0.558	0.483	0.556	0.995	0.995	0.996	0.995	0.994	0.994	0.995	0.997	0.597	0.496
Struc	0.307	0.311	0.366	0.355	0.352	0.353	0.321	0.365	0.997	0.997	0.996	0.997	0.997	0.997	0.997	0.997	0.361	0.304
Pres	0.336	0.334	0.356	0.352	0.355	0.357	0.349	0.367	0.996	0.996	0.997	0.998	0.997	0.997	0.996	0.997	0.34	0.295
Context	0.32	0.318	0.349	0.333	0.343	0.325	0.337	0.594	0.592	0.594	0.596	0.588	0.592	0.591	0.592	0.334	0.289	
PV_R	0.352	0.352	0.341	0.327	0.329	0.438	0.289	0.353	1	1	1	1	1	1	1	1	0.342	0.375
PV_C	0.326	0.331	0.431	0.429	0.431	0.421	0.357	0.433	0.999	1	0.998	1	0.999	1	0.999	1	0.432	0.404
PV_GlR	0.315	0.303	0.363	0.355	0.355	0.368	0.322	0.363	1	1	1	1	1	1	1	1	0.374	0.317
PV_GlC	0.282	0.281	0.302	0.292	0.304	0.288	0.311	1	1	1	1	1	1	1	1	1	0.295	0.32
Ngrams	0.368	0.376	0.474	0.479	0.469	0.475	0.397	0.466	0.995	0.996	0.997	0.997	0.997	0.996	0.996	0.996	0.496	0.452
SPIRIT																		
BOWSrf	0.646	0.646	0.634	0.634	0.639	0.639	0.661	0.644	0.968	0.973	0.965	0.97	0.969	0.968	0.968	0.962	0.67	0.643
TFIDFSrf	0.652	0.656	0.637	0.644	0.644	0.651	0.648	0.646	0.973	0.977	0.969	0.976	0.976	0.974	0.971	0.97	0.65	0.646
Struc	0.635	0.624	0.643	0.638	0.63	0.637	0.633	0.638	0.955	0.96	0.96	0.959	0.964	0.958	0.959	0.954	0.649	0.622
Pres	0.589	0.598	0.591	0.597	0.59	0.589	0.6	0.598	0.951	0.951	0.949	0.943	0.949	0.953	0.946	0.947	0.581	0.585
Context	0.55	0.545	0.544	0.546	0.539	0.542	0.547	0.55	0.679	0.683	0.676	0.682	0.679	0.681	0.671	0.679	0.531	0.545
PV_R	0.622	0.634	0.598	0.608	0.627	0.606	0.613	0.601	0.976	0.972	0.971	0.973	0.972	0.979	0.972	0.971	0.6	0.608
PV_C	0.596	0.601	0.603	0.617	0.609	0.606	0.606	0.603	0.979	0.982	0.977	0.979	0.981	0.98	0.976	0.978	0.605	0.595
PV_GlR	0.591	0.575	0.583	0.592	0.59	0.59	0.597	0.583	0.979	0.979	0.978	0.976	0.976	0.974	0.977	0.567	0.578	
PV_GlC	0.583	0.582	0.594	0.594	0.591	0.581	0.589	0.588	0.98	0.98	0.98	0.979	0.979	0.976	0.977	0.587	0.588	
Ngrams	0.603	0.599	0.606	0.612	0.607	0.61	0.593	0.61	0.979	0.978	0.977	0.98	0.978	0.978	0.979	0.973	0.6	0.59

Table 6: The performance of the different machine learning tasks applied on the different features using random forests. The hierarchy construction method are abbreviated as follows: balanced k -means clustering (BkM), predictive clustering tree (PCT), clustering with complete linkage (CL), clustering with single linkage (SL), random (RND) and manual (MAN). The evaluation measure is *subset accuracy*, where larger value means better result.

20-genres	HMC						HSC						MLC	SC				
	CL	SL	B2M	B3M	B4M	PCT	RND	MAN	CL	SL	B2M	B3M	B4M	PCT	RND	MAN		
BOWSrf	0.244	0.233	0.279	0.276	0.285	0.291	0.233	0.294	0.989	0.989	0.991	0.989	0.989	0.989	0.989	0.305	0.257	
TFIDFSrf	0.239	0.244	0.311	0.307	0.311	0.298	0.264	0.305	0.99	0.989	0.988	0.986	0.985	0.986	0.986	0.326	0.272	
Struc	0.119	0.142	0.164	0.16	0.161	0.168	0.135	0.166	0.989	0.991	0.992	0.991	0.991	0.992	0.992	0.18	0.116	
Pres	0.187	0.179	0.197	0.194	0.186	0.198	0.198	0.207	0.987	0.986	0.989	0.991	0.988	0.984	0.988	0.192	0.155	
Context	0.11	0.11	0.127	0.119	0.122	0.126	0.109	0.122	0.405	0.405	0.41	0.412	0.401	0.4	0.404	0.402	0.128	0.168
PV_R	0.096	0.087	0.211	0.193	0.191	0.116	0.165	0.21	0.999	0.998	0.998	0.999	0.999	0.998	0.999	0.999	0.222	0.124
PV_C	0.163	0.167	0.255	0.249	0.258	0.24	0.19	0.246	0.998	0.998	0.998	0.998	0.998	0.999	0.998	0.27	0.187	
PV_GI_R	0.081	0.076	0.104	0.105	0.115	0.077	0.105	0.999	0.999	0.999	0.999	0.999	0.999	0.998	0.999	0.126	0.105	
PV_GI_C	0.13	0.131	0.168	0.165	0.153	0.173	0.14	0.173	0.998	0.999	0.999	0.999	0.999	0.999	0.999	0.17	0.143	
Ngrams	0.197	0.204	0.258	0.264	0.268	0.271	0.218	0.267	0.989	0.99	0.993	0.993	0.993	0.993	0.991	0.992	0.282	0.234
SPIRIT																		
BOWSrf	0.316	0.329	0.335	0.35	0.348	0.343	0.321	0.35	0.909	0.918	0.903	0.918	0.906	0.913	0.909	0.323	0.326	
TFIDFSrf	0.322	0.322	0.35	0.335	0.342	0.346	0.316	0.349	0.921	0.928	0.906	0.923	0.923	0.918	0.913	0.339	0.34	
Struc	0.346	0.332	0.348	0.342	0.34	0.342	0.343	0.322	0.891	0.896	0.882	0.891	0.896	0.894	0.894	0.352	0.316	
Pres	0.329	0.315	0.311	0.308	0.305	0.304	0.304	0.325	0.866	0.867	0.869	0.865	0.865	0.87	0.854	0.305	0.281	
Context	0.226	0.214	0.211	0.216	0.213	0.216	0.207	0.21	0.358	0.355	0.356	0.365	0.35	0.353	0.357	0.207	0.204	
PV_R	0.291	0.282	0.336	0.328	0.295	0.333	0.322	0.316	0.93	0.929	0.933	0.931	0.942	0.929	0.926	0.305	0.287	
PV_C	0.304	0.304	0.295	0.321	0.311	0.316	0.292	0.312	0.944	0.943	0.93	0.936	0.948	0.942	0.931	0.934	0.289	
PV_GI_R	0.279	0.267	0.262	0.258	0.271	0.267	0.258	0.262	0.951	0.948	0.945	0.938	0.943	0.94	0.942	0.252	0.268	
PV_GI_C	0.289	0.279	0.301	0.294	0.289	0.289	0.291	0.94	0.94	0.937	0.939	0.947	0.94	0.935	0.939	0.288	0.291	
Ngrams	0.316	0.305	0.299	0.304	0.301	0.296	0.292	0.318	0.935	0.935	0.928	0.935	0.93	0.943	0.931	0.921	0.305	0.315

Table 7: The performance of the different machine learning tasks applied on the different features using random forests. The hierarchy construction method are abbreviated as follows: balanced k -means clustering (BkM), predictive clustering tree (PCT), clustering with complete linkage (CL), clustering with single linkage (SL) and manual (MAN). The evaluation measure is *micro precision* where larger value means better result.

20-genres	HMC										HSC				MLC		SC	
	CL	SL	B2M	B3M	B4M	PCT	RND	MAN	CL	SL	B2M	B3M	B4M	PCT	RND	MAN		
BOWSrf	0.51	0.5	0.523	0.525	0.526	0.526	0.498	0.535	0.997	0.997	0.998	0.994	0.998	0.998	0.992	0.533	0.527	
TFIDFSrf	0.512	0.513	0.55	0.543	0.555	0.545	0.516	0.545	0.997	0.997	0.997	0.997	0.996	0.997	0.997	0.55	0.578	
Struc	0.356	0.357	0.405	0.395	0.399	0.39	0.373	0.406	0.993	0.994	0.993	0.994	0.994	0.993	0.994	0.402	0.322	
Pres	0.397	0.396	0.405	0.4	0.403	0.399	0.395	0.417	0.991	0.991	0.993	0.993	0.993	0.991	0.991	0.4	0.387	
Context	0.293	0.3	0.291	0.298	0.289	0.293	0.691	0.694	0.695	0.692	0.692	0.694	0.696	0.701	0.291	0.386		
PV_R	0.308	0.305	0.495	0.501	0.49	0.368	0.429	0.498	0.999	0.999	0.999	1	1	1	0.999	0.999	0.364	
PV_C	0.404	0.415	0.507	0.502	0.512	0.496	0.439	0.502	1	0.999	1	1	1	1	0.999	0.524	0.471	
PV_GlR	0.284	0.279	0.327	0.322	0.331	0.335	0.289	0.333	0.999	0.999	0.999	1	0.999	0.999	1	0.336	0.32	
PV_GlC	0.367	0.384	0.435	0.441	0.428	0.445	0.387	0.45	0.999	0.999	0.999	1	1	1	1	0.466	0.419	
Ngrams	0.435	0.449	0.503	0.502	0.503	0.498	0.454	0.491	0.997	0.997	0.999	0.999	0.999	0.999	0.999	0.498	0.511	
SPIRIT																		
BOWSrf	0.597	0.6	0.626	0.643	0.638	0.646	0.598	0.642	0.965	0.967	0.965	0.968	0.967	0.969	0.963	0.972	0.604	0.613
TFIDFSrf	0.594	0.601	0.649	0.644	0.647	0.652	0.586	0.649	0.973	0.973	0.967	0.972	0.973	0.973	0.974	0.972	0.647	0.612
Struc	0.623	0.611	0.625	0.627	0.619	0.616	0.618	0.613	0.96	0.962	0.954	0.96	0.961	0.962	0.962	0.959	0.638	0.624
Pres	0.592	0.593	0.58	0.584	0.576	0.583	0.581	0.584	0.953	0.953	0.954	0.953	0.953	0.955	0.956	0.948	0.957	0.575
Context	0.512	0.504	0.504	0.507	0.5	0.505	0.506	0.636	0.64	0.637	0.644	0.639	0.643	0.637	0.64	0.496	0.504	
PV_R	0.559	0.564	0.613	0.607	0.56	0.626	0.618	0.605	0.984	0.986	0.983	0.986	0.986	0.987	0.982	0.983	0.602	0.584
PV_C	0.581	0.591	0.578	0.599	0.596	0.588	0.581	0.584	0.986	0.986	0.982	0.986	0.983	0.98	0.982	0.579	0.596	
PV_GlR	0.539	0.527	0.529	0.53	0.532	0.533	0.535	0.525	0.992	0.992	0.989	0.99	0.994	0.99	0.988	0.994	0.521	0.563
PV_GlC	0.579	0.573	0.585	0.581	0.578	0.569	0.568	0.579	0.982	0.982	0.986	0.984	0.986	0.985	0.986	0.987	0.584	0.581
Ngrams	0.587	0.589	0.579	0.586	0.581	0.584	0.574	0.586	0.972	0.974	0.971	0.972	0.973	0.978	0.971	0.972	0.585	0.615

Table 8: The performance of the different machine learning tasks applied on the different features using random forests. The hierarchy construction method are abbreviated as follows: balanced k -means clustering (BkM), predictive clustering tree (PCT), clustering with complete linkage (CL), clustering with single linkage (SL), random (RND) and manual (MAN). The evaluation measure is *micro recall* where larger value means better result.

20-genres	HMC						HSC						MLC		SC			
	CL	SL	B2M	B3M	B4M	PCT	RND	MAN	CL	SL	B2M	B3M	B4M	PCT	RND	MAN		
BOWSrf	0.433	0.417	0.53	0.534	0.52	0.529	0.437	0.538	0.994	0.994	0.999	0.993	0.993	0.999	0.999	0.57	0.512	
TFIDFSrf	0.456	0.453	0.574	0.564	0.568	0.568	0.474	0.564	0.995	0.993	0.994	0.993	0.992	0.992	0.618	0.497		
Struc	0.305	0.301	0.355	0.347	0.342	0.341	0.317	0.354	0.999	0.999	0.998	0.999	0.999	0.999	0.999	0.348	0.326	
Pres	0.326	0.324	0.348	0.34	0.348	0.344	0.335	0.358	0.999	0.999	0.999	0.999	1	0.999	0.999	0.329	0.292	
Context	0.348	0.345	0.383	0.366	0.368	0.37	0.356	0.369	0.595	0.593	0.595	0.596	0.592	0.588	0.585	0.591	0.364	0.278
PV_R	0.411	0.408	0.314	0.304	0.306	0.509	0.265	0.324	1	1	1	1	1	1	1	0.31	0.437	
PV_C	0.319	0.321	0.413	0.414	0.42	0.412	0.344	0.418	0.999	1	0.999	1	0.999	1	0.999	0.417	0.421	
PV_GI_R	0.361	0.352	0.415	0.401	0.408	0.414	0.361	0.41	1	1	1	1	1	1	1	0.416	0.362	
PV_GI_C	0.275	0.278	0.292	0.293	0.282	0.294	0.279	0.298	1	1	1	1	1	1	1	0.285	0.32	
Ngrams	0.36	0.364	0.47	0.479	0.462	0.47	0.383	0.461	0.993	0.995	0.995	0.996	0.996	0.995	0.995	0.493	0.457	
SPIRIT																		
BOWSrf	0.633	0.63	0.587	0.583	0.595	0.588	0.649	0.594	0.974	0.977	0.967	0.975	0.97	0.972	0.973	0.665	0.633	
TFIDFSrf	0.638	0.647	0.588	0.59	0.601	0.602	0.641	0.594	0.973	0.978	0.971	0.976	0.977	0.975	0.971	0.592	0.64	
Struc	0.591	0.58	0.6	0.587	0.591	0.598	0.59	0.599	0.96	0.965	0.965	0.963	0.968	0.961	0.959	0.957	0.587	
Pres	0.519	0.535	0.53	0.545	0.538	0.531	0.545	0.549	0.955	0.956	0.954	0.948	0.953	0.956	0.951	0.949	0.521	0.544
Context	0.504	0.504	0.503	0.496	0.503	0.495	0.503	0.51	0.665	0.673	0.658	0.671	0.666	0.653	0.668	0.482	0.507	
PV_R	0.619	0.625	0.546	0.551	0.62	0.557	0.549	0.971	0.969	0.971	0.97	0.975	0.975	0.969	0.968	0.542	0.603	
PV_C	0.558	0.555	0.563	0.586	0.576	0.572	0.575	0.573	0.977	0.977	0.974	0.977	0.979	0.974	0.975	0.559	0.572	
PV_GI_R	0.554	0.533	0.559	0.555	0.555	0.55	0.573	0.543	0.976	0.973	0.975	0.969	0.97	0.971	0.968	0.517	0.553	
PV_GI_C	0.527	0.529	0.545	0.538	0.533	0.529	0.546	0.54	0.979	0.978	0.973	0.976	0.981	0.972	0.973	0.523	0.555	
Ngrams	0.57	0.568	0.582	0.584	0.582	0.583	0.561	0.581	0.985	0.983	0.982	0.984	0.982	0.984	0.982	0.977	0.564	0.557

Table 9: The performance of the different machine learning tasks applied on the different features using random forests. The hierarchy construction method are abbreviated as follows: balanced k -means clustering (BkM), predictive clustering tree (PCT), clustering with complete linkage (CL), clustering with single linkage (SL), random (RND) and manual (MAN). The evaluation measure is *micro F₁* where larger value means better result.

20-genres	HMC						HSC						MLC		SC			
	CL	SL	B2M	B3M	B4M	PCT	RND	MAN	CL	SL	B2M	B3M	B4M	PCT	RND	MAN		
BOWSrf	0.468	0.455	0.526	0.529	0.523	0.528	0.466	0.537	0.996	0.995	0.996	0.995	0.996	0.995	0.996	0.996	0.551	0.519
TFIDFSrf	0.483	0.481	0.562	0.553	0.562	0.556	0.494	0.555	0.996	0.995	0.995	0.994	0.995	0.994	0.994	0.994	0.582	0.535
Struc	0.329	0.327	0.379	0.369	0.368	0.364	0.343	0.378	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.373	0.324
Pres	0.358	0.356	0.374	0.368	0.373	0.369	0.362	0.385	0.995	0.995	0.996	0.996	0.996	0.995	0.994	0.995	0.362	0.333
Context	0.318	0.317	0.336	0.324	0.325	0.33	0.319	0.327	0.64	0.64	0.642	0.641	0.637	0.636	0.636	0.641	0.324	0.323
PV_R	0.352	0.349	0.384	0.379	0.376	0.427	0.328	0.392	1	0.999	0.999	1	1	1	0.999	1	0.391	0.397
PV_C	0.356	0.362	0.455	0.454	0.461	0.45	0.386	0.456	0.999	0.999	1	0.999	0.999	1	0.999	1	0.465	0.445
PV_GLR	0.318	0.311	0.366	0.357	0.365	0.37	0.321	0.367	1	1	1	1	1	1	0.999	1	0.372	0.34
PV_GLC	0.315	0.322	0.349	0.352	0.34	0.354	0.325	0.359	0.999	1	1	1	1	1	0.999	1	0.354	0.363
Ngrams	0.394	0.402	0.486	0.49	0.481	0.484	0.415	0.476	0.995	0.996	0.997	0.997	0.997	0.996	0.997	0.997	0.495	0.483
SPIRIT																		
BOWSrf	0.614	0.615	0.606	0.612	0.616	0.615	0.622	0.617	0.969	0.972	0.966	0.971	0.968	0.968	0.968	0.633	0.623	
TFIDFSrf	0.615	0.623	0.617	0.616	0.623	0.626	0.612	0.62	0.973	0.975	0.969	0.974	0.975	0.974	0.972	0.97	0.618	0.626
Struc	0.606	0.595	0.612	0.606	0.605	0.607	0.604	0.606	0.96	0.964	0.958	0.962	0.964	0.961	0.961	0.958	0.618	0.605
Pres	0.553	0.562	0.554	0.564	0.556	0.556	0.562	0.566	0.954	0.955	0.954	0.951	0.954	0.954	0.956	0.95	0.553	0.547
Context	0.508	0.504	0.504	0.505	0.498	0.498	0.504	0.508	0.651	0.656	0.647	0.657	0.652	0.653	0.645	0.653	0.489	0.506
PV_R	0.587	0.593	0.578	0.578	0.589	0.589	0.586	0.576	0.978	0.977	0.977	0.978	0.978	0.981	0.975	0.975	0.593	0.593
PV_C	0.569	0.573	0.57	0.593	0.586	0.58	0.578	0.578	0.982	0.981	0.977	0.979	0.983	0.981	0.977	0.978	0.569	0.584
PV_GLR	0.546	0.53	0.544	0.542	0.543	0.542	0.554	0.534	0.984	0.983	0.982	0.979	0.982	0.98	0.979	0.981	0.519	0.558
PV_GLC	0.552	0.55	0.564	0.559	0.554	0.548	0.557	0.559	0.98	0.98	0.98	0.983	0.981	0.979	0.98	0.98	0.552	0.568
Ngrams	0.578	0.579	0.58	0.585	0.581	0.583	0.568	0.583	0.978	0.978	0.976	0.978	0.977	0.981	0.977	0.974	0.575	0.585

Table 10: The performance of the different machine learning tasks applied on the different features using random forests. The hierarchy construction method are abbreviated as follows: balanced k -means clustering (BkM), predictive clustering tree (PCT), clustering with complete linkage (CL), clustering with single linkage (SL) and manual (MAN). The evaluation measure is *macro precision* where larger value means better result.

20-genres	HMC						HSC						MLC	SC				
	CL	SL	B2M	B3M	B4M	PCT	RND	MAN	CL	SL	B2M	B3M	B4M	PCT	RND	MAN		
BOWSrf	0.638	0.65	0.669	0.648	0.623	0.626	0.673	0.677	0.996	0.997	0.998	0.993	0.997	0.998	0.992	0.647	0.571	
TFIDFSrf	0.71	0.644	0.684	0.666	0.649	0.661	0.629	0.629	0.997	0.996	0.997	0.996	0.995	0.997	0.996	0.648	0.639	
Struc	0.424	0.459	0.533	0.502	0.451	0.475	0.496	0.54	0.993	0.994	0.994	0.994	0.994	0.994	0.994	0.528	0.364	
Pres	0.427	0.428	0.424	0.395	0.421	0.398	0.425	0.415	0.991	0.991	0.993	0.993	0.992	0.989	0.991	0.403	0.414	
Context	0.387	0.376	0.39	0.377	0.371	0.39	0.38	0.373	0.808	0.814	0.804	0.808	0.806	0.808	0.811	0.806	0.367	0.432
PV_R	0.549	0.585	0.654	0.764	0.743	0.557	0.535	0.659	0.999	0.998	0.999	0.999	1	0.999	0.999	0.654	0.429	
PV_C	0.532	0.585	0.683	0.652	0.707	0.692	0.731	0.629	1	0.999	1	1	1	1	1	0.637	0.579	
PV_GI_R	0.522	0.519	0.562	0.628	0.603	0.572	0.592	0.592	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.53	0.381	
PV_GI_C	0.339	0.439	0.601	0.615	0.701	0.742	0.601	0.696	0.999	0.999	0.999	0.999	1	1	1	0.711	0.493	
Ngrams	0.636	0.66	0.633	0.678	0.639	0.676	0.609	0.67	0.997	0.997	0.999	0.998	0.998	0.999	0.998	0.632	0.578	
<hr/>																		
SPiRiT																		
BOWSrf	0.329	0.339	0.361	0.34	0.339	0.356	0.335	0.352	0.921	0.993	0.92	0.922	0.994	0.994	0.92	0.994	0.315	
TFIDFSrf	0.352	0.355	0.36	0.352	0.434	0.428	0.318	0.351	0.995	0.923	0.923	0.923	0.923	0.923	0.923	0.41	0.382	
Struc	0.278	0.217	0.371	0.341	0.355	0.341	0.344	0.271	0.992	0.993	0.992	0.993	0.992	0.992	0.992	0.341	0.256	
Pres	0.188	0.259	0.183	0.258	0.253	0.255	0.188	0.26	0.919	0.848	0.919	0.848	0.848	0.918	0.849	0.185	0.179	
Context	0.169	0.152	0.154	0.164	0.156	0.17	0.161	0.616	0.61	0.552	0.625	0.618	0.622	0.549	0.618	0.155	0.16	
PV_R	0.207	0.21	0.232	0.236	0.276	0.305	0.235	0.292	0.925	0.926	0.997	0.997	0.997	0.996	0.997	0.333	0.249	
PV_C	0.227	0.223	0.215	0.242	0.3	0.226	0.226	0.277	0.926	0.997	0.996	0.997	0.997	0.997	0.997	0.925	0.272	
PV_GI_R	0.223	0.22	0.21	0.21	0.212	0.221	0.22	0.218	0.927	0.998	0.927	0.999	0.998	0.998	0.999	0.235	0.249	
PV_GI_C	0.228	0.212	0.228	0.218	0.22	0.212	0.224	0.232	0.997	0.997	0.997	0.997	0.997	0.997	0.996	0.229	0.233	
Ngrams	0.227	0.233	0.286	0.293	0.285	0.227	0.224	0.291	0.994	0.995	0.993	0.994	0.995	0.994	0.995	0.407	0.413	

Table 11: The performance of the different machine learning tasks applied on the different features using random forests. The hierarchy construction method are abbreviated as follows: balanced k -means clustering (BkM), predictive clustering tree (PCT), clustering with complete linkage (CL), clustering with single linkage (SL) and manual (MAN). The evaluation measure is *macro recall* where larger value means better result.

20-genres	HMC										HSC				MLC		SC	
	CL	SL	B2M	B3M	B4M	PCT	RND	MAN	CL	SL	B2M	B3M	B4M	PCT	RND	MAN		
BOWSrf	0.378	0.361	0.501	0.503	0.488	0.503	0.381	0.51	0.992	0.993	0.993	0.999	0.992	0.991	0.992	0.999	0.555	0.518
TFIDFSrf	0.404	0.403	0.552	0.543	0.545	0.548	0.425	0.543	0.993	0.991	0.992	0.991	0.989	0.991	0.99	0.99	0.606	0.517
Struc	0.232	0.23	0.297	0.294	0.285	0.289	0.251	0.302	0.998	0.998	0.997	0.998	0.998	0.998	0.998	0.998	0.309	0.294
Pres	0.267	0.262	0.301	0.293	0.298	0.3	0.279	0.312	0.999	0.998	0.999	0.999	0.999	0.999	0.999	0.999	0.293	0.263
Context	0.285	0.28	0.323	0.306	0.307	0.307	0.292	0.308	0.537	0.535	0.534	0.535	0.529	0.526	0.532	0.532	0.307	0.251
PV_R	0.304	0.301	0.278	0.266	0.264	0.451	0.191	0.282	1	1	1	1	1	1	1	1	0.295	0.4
PV_C	0.228	0.232	0.376	0.371	0.379	0.38	0.269	0.384	0.999	1	0.998	1	0.998	0.999	0.999	0.998	0.402	0.403
PV_GlR	0.241	0.234	0.332	0.318	0.324	0.333	0.254	0.332	1	1	1	1	1	1	1	1	0.356	0.311
PV_GlC	0.183	0.186	0.235	0.227	0.219	0.235	0.19	0.24	1	1	1	1	1	0.998	0.999	0.997	0.25	0.285
Ngrams	0.301	0.3	0.448	0.452	0.436	0.443	0.33	0.438	0.992	0.994	0.994	0.995	0.995	0.994	0.994	0.994	0.465	
SPIRIT																		
BOWSrf	0.23	0.23	0.207	0.207	0.214	0.212	0.237	0.215	0.764	0.778	0.721	0.798	0.739	0.777	0.737	0.701	0.272	0.248
TFIDFSrf	0.233	0.235	0.211	0.21	0.221	0.222	0.235	0.219	0.743	0.755	0.725	0.754	0.739	0.756	0.733	0.714	0.238	0.255
Struc	0.211	0.206	0.219	0.215	0.216	0.221	0.213	0.221	0.714	0.705	0.718	0.716	0.756	0.716	0.703	0.677	0.225	0.218
Pres	0.174	0.181	0.179	0.188	0.185	0.181	0.185	0.191	0.646	0.639	0.65	0.639	0.629	0.638	0.635	0.601	0.177	0.188
Context	0.147	0.148	0.149	0.148	0.144	0.143	0.147	0.15	0.287	0.312	0.283	0.301	0.291	0.287	0.285	0.301	0.139	0.151
PV_R	0.214	0.216	0.183	0.183	0.219	0.187	0.183	0.18	0.723	0.69	0.717	0.719	0.716	0.748	0.711	0.699	0.184	0.218
PV_C	0.183	0.178	0.186	0.196	0.193	0.187	0.188	0.195	0.761	0.738	0.751	0.766	0.785	0.759	0.746	0.724	0.188	0.2
PV_GlR	0.173	0.159	0.177	0.174	0.173	0.173	0.182	0.167	0.748	0.728	0.733	0.711	0.715	0.693	0.719	0.696	0.161	0.181
PV_GlC	0.16	0.16	0.172	0.168	0.166	0.17	0.171	0.771	0.765	0.756	0.742	0.764	0.76	0.725	0.726	0.163	0.182	
Ngrams	0.193	0.192	0.202	0.203	0.202	0.203	0.188	0.205	0.869	0.841	0.846	0.845	0.827	0.838	0.835	0.771	0.202	0.2

Table 12: The performance of the different machine learning tasks applied on the different features using random forests. The hierarchy construction method are abbreviated as follows: balanced k -means clustering (BkM), predictive clustering tree (PCT), clustering with complete linkage (CL), clustering with single linkage (SL) and manual (MAN). The evaluation measure is *macro F₁* where larger value means better result.

20-genres	HMC						HSC						MLC	SC				
	CL	SL	B2M	B3M	B4M	PCT	RND	MAN	CL	SL	B2M	B3M	B4M	PCT	RND	MAN		
BOWSrf	0.424	0.409	0.529	0.53	0.52	0.528	0.418	0.542	0.994	0.995	0.996	0.995	0.994	0.995	0.996	0.572	0.531	
TFIDFSrf	0.454	0.448	0.569	0.566	0.566	0.568	0.463	0.563	0.995	0.994	0.994	0.993	0.993	0.993	0.993	0.605	0.563	
Struc	0.247	0.248	0.315	0.315	0.304	0.304	0.267	0.328	0.995	0.996	0.996	0.996	0.996	0.996	0.996	0.331	0.31	
Pres	0.296	0.289	0.324	0.316	0.324	0.321	0.304	0.337	0.995	0.995	0.996	0.996	0.996	0.994	0.995	0.318	0.299	
Context	0.287	0.28	0.318	0.304	0.302	0.307	0.291	0.304	0.614	0.615	0.611	0.609	0.606	0.604	0.611	0.302	0.295	
PV_R	0.29	0.286	0.347	0.341	0.332	0.444	0.219	0.35	1	0.999	0.999	1	1	0.999	0.999	0.374	0.391	
PV_C	0.242	0.253	0.436	0.431	0.437	0.442	0.3	0.439	0.999	0.999	0.999	0.999	0.999	1	0.999	0.462	0.451	
PV_GI_R	0.208	0.199	0.335	0.33	0.337	0.346	0.237	0.347	1	0.99	0.99	0.999	0.999	1	1	0.999	0.373	
PV_GI_C	0.17	0.179	0.268	0.26	0.25	0.276	0.189	0.285	0.999	0.999	1	1	0.999	1	0.999	0.311	0.338	
Ngrams	0.331	0.333	0.486	0.492	0.48	0.484	0.363	0.482	0.994	0.995	0.996	0.997	0.997	0.996	0.996	0.515	0.497	
SPIRIT																		
BOWSrf	0.223	0.221	0.21	0.216	0.22	0.223	0.224	0.811	0.837	0.783	0.85	0.804	0.837	0.785	0.768	0.273	0.251	
TFIDFSrf	0.219	0.221	0.222	0.217	0.234	0.232	0.223	0.231	0.804	0.803	0.784	0.805	0.783	0.81	0.789	0.771	0.266	0.261
Struc	0.21	0.202	0.222	0.223	0.221	0.223	0.213	0.223	0.786	0.771	0.788	0.791	0.826	0.783	0.772	0.744	0.233	0.222
Pres	0.175	0.182	0.177	0.188	0.183	0.181	0.182	0.19	0.707	0.696	0.709	0.713	0.683	0.69	0.7	0.662	0.175	0.181
Context	0.136	0.134	0.136	0.137	0.133	0.131	0.136	0.137	0.309	0.334	0.308	0.328	0.319	0.314	0.306	0.326	0.13	0.141
PV_R	0.199	0.2	0.192	0.19	0.205	0.197	0.19	0.187	0.783	0.744	0.778	0.787	0.785	0.814	0.777	0.761	0.2	0.218
PV_C	0.185	0.178	0.183	0.194	0.194	0.185	0.183	0.196	0.811	0.793	0.813	0.83	0.85	0.811	0.796	0.769	0.194	0.203
PV_GI_R	0.168	0.147	0.117	0.166	0.165	0.168	0.171	0.159	0.805	0.797	0.787	0.779	0.784	0.749	0.783	0.761	0.164	0.185
PV_GI_C	0.161	0.159	0.172	0.169	0.171	0.168	0.166	0.174	0.827	0.823	0.818	0.793	0.812	0.825	0.781	0.781	0.17	0.182
Ngrams	0.193	0.192	0.2	0.201	0.199	0.201	0.186	0.205	0.92	0.887	0.898	0.896	0.884	0.892	0.883	0.829	0.212	0.216

Table 13: The performance of the different machine learning tasks applied on the different features using random forests. The hierarchy construction method are abbreviated as follows: balanced k -means clustering (BkM), predictive clustering tree (PCT), clustering with complete linkage (CL), clustering with single linkage (SL), random (RND) and manual (MAN). The evaluation measure is *one-error* where smaller value means better result.

20-genres	HMC										HSC				MLC		SC	
	CL	SL	B2M	B3M	B4M	PCT	RND	MAN	CL	SL	B2M	B3M	B4M	PCT	RND	MAN		
BOWSrf	0.488	0.503	0.435	0.42	0.436	0.439	0.481	0.414	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.402	0.362	
TFIDFSrf	0.468	0.467	0.385	0.402	0.396	0.4	0.45	0.407	0.002	0.001	0.001	0.001	0.001	0.001	0.001	0.368	0.361	
Struc	0.625	0.641	0.595	0.591	0.583	0.599	0.634	0.582	0.002	0.002	0.001	0.002	0.002	0.002	0.002	0.59	0.619	
Pres	0.612	0.608	0.593	0.598	0.593	0.591	0.608	0.591	0.001	0	0	0.001	0.001	0	0.002	0.001	0.6	
Context	0.661	0.662	0.643	0.641	0.644	0.638	0.652	0.648	0.326	0.317	0.321	0.32	0.319	0.32	0.32	0.639	0.633	
PV_R	0.64	0.643	0.555	0.561	0.556	0.55	0.607	0.561	0	0	0	0	0	0	0	0.531	0.535	
PV_C	0.596	0.59	0.484	0.48	0.479	0.486	0.558	0.494	0	0	0	0	0	0	0	0.48	0.467	
PV_GLR	0.664	0.671	0.629	0.622	0.628	0.618	0.664	0.618	0	0	0	0	0	0	0	0.61	0.589	
PV_GlC	0.635	0.637	0.598	0.606	0.611	0.592	0.626	0.593	0	0	0	0	0	0	0	0.584	0.531	
Ngrams	0.568	0.555	0.465	0.469	0.474	0.47	0.546	0.48	0.002	0	0	0	0	0.001	0.001	0	0.453	0.405
SPIRIT																		
BOWSrf	0.313	0.322	0.308	0.305	0.313	0.318	0.311	0.016	0.016	0.013	0.017	0.014	0.017	0.018	0.019	0.308	0.302	
TFIDFSrf	0.312	0.311	0.312	0.313	0.305	0.313	0.311	0.309	0.011	0.011	0.011	0.011	0.007	0.01	0.012	0.301	0.306	
Struc	0.325	0.326	0.316	0.318	0.319	0.33	0.335	0.313	0.035	0.034	0.029	0.033	0.037	0.035	0.032	0.304	0.309	
Pres	0.36	0.36	0.356	0.369	0.352	0.363	0.359	0.366	0.026	0.026	0.022	0.027	0.026	0.026	0.022	0.373	0.359	
Context	0.424	0.426	0.431	0.427	0.426	0.418	0.428	0.427	0.335	0.338	0.341	0.337	0.332	0.337	0.335	0.338	0.444	
PV_R	0.346	0.335	0.352	0.356	0.36	0.342	0.349	0.339	0.003	0.001	0.008	0.006	0.004	0.003	0.008	0.005	0.349	
PV_C	0.343	0.35	0.356	0.35	0.353	0.35	0.343	0.359	0.001	0.003	0.004	0.004	0.001	0.004	0.005	0.355	0.35	
PV_GLR	0.389	0.391	0.396	0.39	0.394	0.389	0.393	0.002	0.004	0.004	0.002	0.003	0.003	0.004	0.002	0.4	0.373	
PV_GlC	0.367	0.36	0.357	0.352	0.372	0.365	0.369	0.366	0.004	0.004	0.002	0.004	0.003	0.002	0.003	0.005	0.377	
Ngrams	0.359	0.365	0.359	0.356	0.357	0.365	0.356	0.006	0.007	0.009	0.007	0.007	0.007	0.012	0.365	0.357		

Table 14: The performance of the different machine learning tasks applied on the different features using random forests. The hierarchy construction method are abbreviated as follows: balanced k -means clustering (BkM), predictive clustering tree (PCT), clustering with complete linkage (CL), clustering with single linkage (SL), random (RND) and manual (MAN). The evaluation measure is coverage where smaller value means better result.

20-genres	HMC						HSC						MLC	SC			
	CL	SL	B2M	B3M	B4M	PCT	RND	MAN	CL	SL	B2M	B3M	B4M	PCT	RND	MAN	
BOWSrf	2.832	2.816	2.424	2.368	2.414	2.427	2.849	2.395	0.342	0.342	0.342	0.342	0.342	0.342	0.342	2.322	2.19
TFIDFSrf	2.641	2.761	2.148	2.196	2.146	2.195	2.604	2.194	0.343	0.342	0.342	0.342	0.342	0.342	0.342	2.096	2.175
Struc	4.157	4.29	3.873	3.867	3.859	3.814	4.018	3.841	0.344	0.344	0.343	0.344	0.344	0.344	0.343	3.86	4.504
Pres	4.726	4.551	4.361	4.505	4.501	4.397	4.604	4.603	0.343	0.341	0.342	0.342	0.341	0.343	0.342	4.542	4.41
Context	5.211	5.185	5.006	5.02	4.957	4.96	5.088	5.045	1.168	1.179	1.167	1.174	1.176	1.178	1.182	5.158	5.201
PV_R	3.986	4.008	3.177	3.217	3.332	3.199	3.665	3.316	0.342	0.342	0.342	0.342	0.342	0.342	0.342	3.193	3.344
PV_C	3.483	3.42	2.714	2.702	2.74	2.719	3.222	2.727	0.342	0.342	0.342	0.342	0.342	0.342	0.342	2.693	2.648
PV_GI_R	4.778	4.863	3.904	3.984	3.935	3.907	4.415	3.852	0.342	0.342	0.342	0.342	0.342	0.342	0.342	3.973	4.203
PV_GI_C	4.292	4.128	3.436	3.483	3.523	3.432	3.944	3.514	0.342	0.342	0.342	0.342	0.342	0.342	0.342	3.462	3.276
Ngrams	3.364	3.309	2.682	2.712	2.85	2.775	3.209	2.769	0.344	0.343	0.342	0.343	0.343	0.343	0.342	2.695	2.389
<hr/>																	
SPiRiT																	
BOWSrf	2.007	2.016	1.94	1.948	1.923	1.932	1.919	1.939	0.59	0.59	0.591	0.594	0.593	0.594	0.596	1.867	1.879
TFIDFSrf	1.977	1.96	1.976	1.915	1.918	1.911	2.001	1.851	0.585	0.587	0.585	0.586	0.588	0.586	0.586	1.843	1.955
Struc	2.135	2.112	2.041	2.07	2.018	2.04	2.104	2.07	0.611	0.613	0.607	0.611	0.614	0.611	0.611	2.043	2.081
Pres	2.238	2.21	2.206	2.211	2.206	2.16	2.213	2.21	0.604	0.604	0.601	0.606	0.602	0.604	0.599	0.604	2.26
Context	2.42	2.396	2.413	2.43	2.406	2.396	2.423	2.399	1.394	1.382	1.384	1.383	1.379	1.363	1.378	2.525	2.513
PV_R	2.186	2.176	2.248	2.204	2.211	2.174	2.226	2.166	0.579	0.576	0.582	0.577	0.577	0.583	0.579	2.166	2.194
PV_C	2.115	2.17	2.067	2.133	2.094	2.122	2.094	2.104	0.575	0.575	0.578	0.577	0.579	0.579	0.58	0.58	2.095
PV_GI_R	2.401	2.438	2.36	2.363	2.357	2.316	2.343	2.359	0.576	0.577	0.576	0.577	0.578	0.577	0.577	2.333	2.302
PV_GI_C	2.217	2.26	2.214	2.204	2.167	2.193	2.221	2.156	0.578	0.579	0.577	0.576	0.577	0.576	0.58	2.183	2.163
Ngrams	2.17	2.201	2.157	2.18	2.156	2.152	2.176	2.152	0.582	0.583	0.584	0.582	0.582	0.582	0.589	2.111	2.034

Table 15: The performance of the different machine learning tasks applied on the different features using random forests. The hierarchy construction method are abbreviated as follows: balanced k -means clustering (BkM), predictive clustering tree (PCT), clustering with complete linkage (CL), clustering with single linkage (SL), random (RND) and manual (MAN). The evaluation measure is *ranking loss* where smaller value means better result.

<i>20-genres</i>	HMC										HSC				MLC		SC	
	CL	SL	B2M	B3M	B4M	PCT	RND	MAN	CL	SL	B2M	B3M	B4M	PCT	RND	MAN		
BOWSrf	0.11	0.108	0.091	0.088	0.089	0.09	0.109	0.088	0	0	0	0	0	0	0	0.084	0.078	
TFIDFSrf	0.101	0.105	0.077	0.08	0.077	0.079	0.097	0.079	0	0	0	0	0	0	0	0.073	0.077	
Struc	0.17	0.177	0.157	0.157	0.156	0.155	0.165	0.156	0	0	0	0	0	0	0	0.156	0.186	
Pres	0.202	0.193	0.185	0.19	0.191	0.186	0.196	0.194	0	0	0	0	0	0	0	0.193	0.186	
Context	0.227	0.225	0.217	0.218	0.215	0.221	0.219	0.039	0.039	0.039	0.039	0.039	0.039	0.038	0.04	0.225	0.227	
PV_R	0.167	0.168	0.129	0.131	0.137	0.129	0.151	0.136	0	0	0	0	0	0	0	0.128	0.133	
PV_C	0.143	0.14	0.105	0.104	0.107	0.105	0.128	0.105	0	0	0	0	0	0	0	0.103	0.1	
PV_GLR	0.206	0.21	0.162	0.167	0.164	0.163	0.188	0.16	0	0	0	0	0	0	0	0.165	0.174	
PV_GLC	0.181	0.175	0.139	0.141	0.142	0.139	0.164	0.142	0	0	0	0	0	0	0	0.14	0.129	
Ngrams	0.133	0.131	0.101	0.102	0.108	0.105	0.126	0.105	0	0	0	0	0	0	0	0.102	0.087	
SPIRIT																		
BOWSrf	0.085	0.088	0.082	0.082	0.081	0.082	0.081	0.08	0.001	0.001	0.001	0.001	0.001	0.001	0.002	0.076	0.078	
TFIDFSrf	0.084	0.083	0.083	0.08	0.08	0.081	0.085	0.076	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.075	0.081	
Struc	0.095	0.094	0.089	0.09	0.088	0.089	0.093	0.09	0.003	0.003	0.002	0.003	0.003	0.003	0.003	0.087	0.09	
Pres	0.104	0.102	0.103	0.103	0.102	0.1	0.102	0.103	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.103	0.104	
Context	0.116	0.114	0.116	0.116	0.114	0.116	0.115	0.054	0.054	0.054	0.054	0.054	0.054	0.054	0.053	0.12	0.119	
PV_R	0.099	0.097	0.102	0.1	0.101	0.098	0.101	0.097	0	0	0.001	0.001	0	0	0.001	0	0.096	0.096
PV_C	0.095	0.097	0.092	0.094	0.094	0.095	0.093	0.093	0	0	0	0	0	0	0	0.092	0.088	
PV_GLR	0.112	0.115	0.11	0.108	0.108	0.109	0.11	0	0	0	0	0	0	0	0	0.109	0.105	
PV_GLC	0.101	0.103	0.1	0.099	0.098	0.099	0.1	0.097	0	0	0	0	0	0	0	0.099	0.096	
Ngrams	0.097	0.1	0.097	0.098	0.097	0.096	0.096	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.093	0.089	

Table 16: The performance of the different machine learning tasks applied on the different features using random forests. The hierarchy construction method are abbreviated as follows: balanced k -means clustering (BkM), predictive clustering tree (PCT), clustering with complete linkage (CL), clustering with single linkage (SL), random (RND) and manual (MAN). The evaluation measure is average precision where larger value means better result.

20-genres	HMC						HSC						MLC		SC	
	CL	SL	B2M	B3M	B4M	PCT	RND	MAN	CL	SL	B2M	B3M	B4M	PCT	RND	MAN
BOWSrf	0.64	0.631	0.682	0.693	0.685	0.682	0.644	0.694	1	0.99	1	1	0.999	1	0.702	0.722
TFIDFSrf	0.656	0.657	0.716	0.706	0.713	0.71	0.669	0.706	0.999	1	1	1	0.999	1	0.727	0.729
Struc	0.519	0.509	0.549	0.549	0.551	0.548	0.523	0.554	0.999	0.999	0.999	0.999	0.999	0.999	0.551	0.519
Pres	0.516	0.517	0.533	0.528	0.528	0.532	0.521	0.533	0.999	1	1	1	1	1	0.999	0.524
Context	0.467	0.469	0.482	0.484	0.483	0.488	0.475	0.48	0.794	0.795	0.798	0.795	0.797	0.797	0.795	0.478
PV_R	0.514	0.51	0.591	0.587	0.584	0.593	0.545	0.59	1	1	1	1	1	1	1	0.605
PV_C	0.553	0.556	0.645	0.644	0.643	0.641	0.586	0.64	1	1	1	1	1	1	1	0.645
PV_GLR	0.474	0.47	0.526	0.525	0.527	0.534	0.489	0.535	1	1	1	1	1	1	1	0.537
PV_GLC	0.503	0.506	0.559	0.551	0.547	0.561	0.521	0.558	1	1	1	1	1	1	1	0.543
Ngrams	0.578	0.589	0.654	0.653	0.647	0.65	0.596	0.646	0.999	1	1	1	1	1	1	0.561
SPIRIT																
BOWSrf	0.764	0.756	0.767	0.77	0.765	0.764	0.765	0.766	0.992	0.992	0.993	0.991	0.992	0.991	0.99	0.774
TFIDFSrf	0.764	0.764	0.767	0.768	0.771	0.768	0.764	0.771	0.994	0.994	0.996	0.995	0.994	0.994	0.778	0.776
Struc	0.749	0.748	0.758	0.757	0.757	0.751	0.746	0.758	0.982	0.982	0.983	0.981	0.982	0.983	0.76	0.765
Pres	0.721	0.727	0.728	0.725	0.728	0.724	0.726	0.724	0.986	0.987	0.986	0.987	0.987	0.989	0.986	0.723
Context	0.689	0.689	0.687	0.687	0.688	0.692	0.685	0.688	0.789	0.789	0.787	0.789	0.789	0.789	0.674	0.682
PV_R	0.736	0.74	0.73	0.731	0.725	0.737	0.732	0.735	0.998	0.999	0.996	0.997	0.998	0.999	0.997	0.736
PV_C	0.737	0.734	0.735	0.739	0.736	0.734	0.737	0.733	0.999	0.999	0.998	0.998	0.998	0.997	0.997	0.744
PV_GLR	0.703	0.7	0.707	0.703	0.705	0.705	0.706	0.706	0.998	0.998	0.999	0.998	0.998	0.999	0.999	0.748
PV_GLC	0.726	0.721	0.726	0.731	0.721	0.725	0.721	0.724	0.998	0.998	0.999	0.999	0.999	0.999	0.997	0.736
Ngrams	0.731	0.727	0.731	0.731	0.734	0.732	0.733	0.736	0.997	0.996	0.995	0.996	0.996	0.996	0.993	0.743