

Appendix 7: Model selection for SEMANTIC-FEATURES (EXP6) in Ch. 10

Gunn Inger Lyse

19th April 2011

Content: Results from model selection (cross-validation) in Chapter 10 with knowledge source=SEMANTIC-FEATURES. The target words are ordered alphabetically. Evaluated with 5-fold cross validation and Overall Accuracy (measured as total recall). The best accuracy in each group is marked in bold-face (in case of ties, the model with the smallest context window is selected).

friskAJ cross-validation results (baseline: 0.683)
SF-SF

100	.427	.427	.439	.451	.415	.439	.390	.390	.366
75	.476	.488	.451	.463	.463	.488	.451	.415	.390
50	.476	.488	.488	.427	.463	.427	.463	.427	.427
30	.537	.537	.549	.488	.488	.488	.500	.427	.402
20	.524	.537	.549	.488	.512	.476	.451	.439	.451
10	.585	.573	.585	.585	.537	.512	.451	.500	.463
4	.610	.622	.537	.512	.573	.524	.463	.488	.451
2	.622	.585	.524	.512	.524	.500	.451	.512	.427
1	.500	.537	.537	.537	.585	.524	.476	.476	.427
	1	2	4	10	20	30	50	75	100

Table 1: friskAJ

fullAJ cross-validation results (baseline: 0.941)
SF-SF

100	.114	.109	.114	.107	.093	.093	.095	.080	.086
75	.107	.105	.107	.116	.091	.091	.091	.082	.091
50	.170	.168	.161	.157	.109	.107	.109	.095	.107
30	.236	.223	.218	.173	.109	.100	.114	.095	.098
20	.420	.386	.341	.239	.136	.118	.109	.098	.100
10	.620	.570	.534	.445	.198	.150	.111	.089	.105
4	.725	.695	.643	.570	.286	.180	.120	.098	.109
2	.709	.677	.577	.559	.316	.177	.127	.100	.107
1	.725	.732	.618	.609	.357	.195	.134	.100	.107
	1	2	4	10	20	30	50	75	100

Table 2: fullAJ

fyrN cross-validation results (baseline: 0.789)

SF-SF									
100	.439	.421	.404	.368	.368	.316	.281	.246	.228
75	.614	.579	.544	.474	.421	.386	.263	.263	.246
50	.754	.754	.754	.719	.614	.544	.421	.263	.228
30	.807	.825	.807	.825	.789	.719	.579	.298	.281
20	.807	.825	.807	.754	.719	.667	.667	.386	.316
10	.754	.789	.754	.754	.596	.719	.649	.421	.316
4	.719	.772	.772	.772	.649	.684	.702	.456	.298
2	.684	.789	.737	.772	.649	.684	.702	.491	.333
1	.614	.737	.649	.737	.596	.667	.684	.456	.316
	1	2	4	10	20	30	50	75	100

Table 3: *fyrN*

galAJ cross-validation results (baseline: 0.776)

SF-SF									
100	.457	.440	.422	.388	.362	.310	.276	.293	.259
75	.448	.448	.440	.405	.422	.379	.328	.293	.250
50	.509	.517	.483	.483	.526	.509	.457	.336	.233
30	.603	.612	.586	.595	.638	.638	.578	.336	.259
20	.621	.603	.621	.569	.621	.612	.543	.379	.284
10	.664	.681	.647	.603	.612	.586	.586	.414	.319
4	.733	.741	.698	.664	.629	.621	.578	.431	.328
2	.690	.672	.655	.664	.690	.638	.603	.440	.345
1	.698	.690	.690	.681	.690	.647	.595	.457	.353
	1	2	4	10	20	30	50	75	100

Table 4: *galAJ*

lagN cross-validation results (baseline: 0.703)

SF-SF									
100	.811	.811	.784	.757	.703	.757	.784	.730	.622
75	.838	.838	.838	.838	.757	.757	.838	.784	.757
50	.838	.838	.838	.811	.811	.811	.757	.757	.703
30	.811	.811	.784	.757	.757	.784	.838	.784	.757
20	.811	.811	.811	.757	.757	.784	.811	.784	.838
10	.865	.838	.838	.811	.784	.811	.838	.784	.811
4	.676	.622	.757	.757	.730	.784	.811	.730	.784
2	.622	.649	.703	.757	.757	.784	.838	.757	.811
1	.649	.676	.676	.757	.811	.757	.865	.784	.811
	1	2	4	10	20	30	50	75	100

Table 5: *lagN*

livN cross-validation results (baseline: 0.981)

SF-SF									
100	.023	.024	.022	.022	.019	.019	.019	.019	.019
75	.022	.023	.023	.021	.019	.019	.021	.019	.019
50	.023	.023	.021	.021	.019	.021	.021	.019	.019
30	.029	.029	.024	.023	.021	.022	.021	.019	.019
20	.124	.117	.065	.033	.024	.022	.021	.019	.019
10	.644	.621	.462	.183	.042	.031	.024	.023	.022
4	.860	.859	.786	.429	.083	.038	.026	.025	.023
2	.887	.915	.846	.532	.109	.039	.028	.026	.024
1	.893	.927	.873	.602	.134	.043	.029	.026	.023
	1	2	4	10	20	30	50	75	100

Table 6: *livN*

planN cross-validation results (baseline: 0.872)

		SF-SF								
100	.514	.495	.495	.486	.505	.495	.450	.450	.294	
75	.642	.651	.615	.642	.606	.624	.550	.514	.422	
50	.642	.642	.633	.642	.615	.596	.615	.532	.468	
30	.688	.697	.661	.661	.587	.569	.596	.541	.459	
20	.716	.725	.697	.651	.615	.541	.569	.514	.459	
10	.789	.798	.752	.780	.697	.596	.569	.495	.413	
4	.789	.798	.789	.798	.697	.587	.560	.486	.450	
2	.771	.807	.697	.752	.670	.596	.560	.495	.468	
1	.697	.761	.734	.771	.706	.624	.578	.495	.459	
	1	2	4	10	20	30	50	75	100	

Table 7: *planN*

rotN cross-validation results (baseline: 0.804)

		SF-SF								
100	.482	.500	.500	.518	.536	.500	.411	.429	.375	
75	.482	.464	.482	.482	.536	.500	.429	.482	.464	
50	.518	.482	.500	.536	.500	.571	.500	.500	.464	
30	.536	.536	.554	.643	.571	.625	.571	.482	.500	
20	.589	.536	.536	.607	.679	.571	.607	.518	.536	
10	.607	.661	.714	.714	.714	.643	.679	.554	.554	
4	.554	.571	.607	.661	.786	.696	.643	.518	.518	
2	.500	.589	.643	.643	.821	.732	.625	.536	.554	
1	.500	.625	.696	.643	.786	.714	.625	.554	.554	
	1	2	4	10	20	30	50	75	100	

Table 8: *rotN*

slagN cross-validation results (baseline: 0.556)

		SF-SF								
100	.233	.233	.226	.233	.248	.226	.211	.188	.188	
75	.233	.226	.248	.241	.195	.248	.180	.188	.150	
50	.241	.248	.203	.211	.226	.241	.218	.211	.150	
30	.323	.308	.323	.263	.233	.226	.226	.211	.143	
20	.338	.316	.346	.301	.293	.241	.226	.188	.158	
10	.376	.383	.368	.338	.368	.338	.278	.278	.203	
4	.316	.331	.383	.323	.383	.338	.286	.316	.211	
2	.368	.406	.398	.429	.451	.391	.331	.308	.195	
1	.346	.406	.429	.421	.444	.391	.346	.301	.203	
	1	2	4	10	20	30	50	75	100	

Table 9: *slagN*

stemmeN cross-validation results (baseline: 0.922)

		SF-SF								
100	.365	.362	.359	.359	.410	.398	.488	.350	.404	
75	.317	.317	.332	.332	.350	.368	.437	.299	.377	
50	.308	.293	.308	.320	.317	.317	.380	.281	.338	
30	.461	.461	.440	.389	.368	.380	.440	.326	.371	
20	.671	.650	.593	.494	.410	.389	.431	.326	.359	
10	.943	.934	.910	.760	.584	.500	.485	.386	.449	
4	.967	.976	.973	.895	.713	.569	.509	.404	.443	
2	.937	.937	.967	.928	.713	.572	.518	.392	.413	
1	.871	.910	.955	.943	.716	.554	.530	.380	.404	
	1	2	4	10	20	30	50	75	100	

Table 10: *stemmeN*

takN cross-validation results (baseline: 0.475)

		SF-SF								
100	.350	.362	.366	.354	.339	.358	.311	.265	.249	
75	.377	.377	.370	.381	.374	.377	.339	.296	.292	
50	.366	.362	.377	.385	.385	.362	.381	.354	.300	
30	.420	.447	.420	.428	.397	.409	.401	.377	.331	
20	.455	.486	.471	.455	.432	.405	.397	.377	.335	
10	.537	.553	.514	.521	.440	.424	.366	.381	.327	
4	.595	.553	.564	.560	.451	.482	.377	.401	.354	
2	.545	.556	.560	.564	.486	.482	.381	.393	.358	
1	.521	.556	.572	.537	.482	.482	.381	.401	.362	
	1	2	4	10	20	30	50	75	100	

Table 11: takN

trykkeV cross-validation results (baseline: 0.804)

		SF-SF								
100	.565	.587	.543	.565	.522	.478	.413	.304	.239	
75	.783	.761	.739	.761	.717	.652	.522	.413	.261	
50	.717	.717	.717	.739	.783	.761	.652	.500	.326	
30	.696	.696	.652	.652	.696	.717	.717	.652	.435	
20	.761	.783	.804	.739	.783	.717	.717	.761	.543	
10	.587	.674	.717	.739	.696	.761	.674	.761	.609	
4	.674	.674	.696	.761	.761	.761	.761	.761	.717	
2	.522	.543	.565	.717	.783	.783	.761	.739	.696	
1	.217	.457	.500	.630	.783	.761	.739	.739	.696	
	1	2	4	10	20	30	50	75	100	

Table 12: trykkeV

utsetteV cross-validation results (baseline: 0.675)

		SF-SF								
100	.519	.519	.558	.506	.532	.532	.545	.532	.519	
75	.494	.494	.481	.442	.506	.494	.532	.519	.532	
50	.494	.468	.519	.468	.481	.442	.506	.532	.532	
30	.584	.571	.571	.584	.506	.494	.571	.506	.571	
20	.636	.597	.610	.688	.584	.558	.532	.519	.584	
10	.610	.584	.584	.649	.584	.519	.468	.494	.519	
4	.714	.662	.571	.662	.688	.532	.416	.481	.519	
2	.597	.610	.545	.701	.675	.558	.468	.468	.519	
1	.636	.623	.532	.662	.675	.597	.455	.481	.519	
	1	2	4	10	20	30	50	75	100	

Table 13: utsetteV

utvalgN cross-validation results (baseline: 0.609)

		SF-SF								
100	.761	.761	.761	.783	.783	.761	.761	.783	.826	
75	.739	.717	.674	.739	.717	.761	.739	.804	.783	
50	.783	.739	.696	.630	.652	.717	.739	.761	.783	
30	.717	.717	.696	.717	.696	.739	.761	.739	.761	
20	.804	.804	.826	.739	.717	.739	.783	.761	.739	
10	.826	.848	.848	.891	.739	.696	.739	.783	.717	
4	.739	.739	.717	.804	.804	.652	.739	.783	.739	
2	.652	.717	.717	.870	.761	.630	.739	.783	.674	
1	.587	.500	.587	.761	.717	.630	.717	.739	.674	
	1	2	4	10	20	30	50	75	100	

Table 14: utvalgN

valgN cross-validation results (baseline: 0.606)

		SF-SF								
	1	2	4	10	20	30	50	75	100	
100	.423	.423	.433	.423	.423	.423	.452	.452	.510	
75	.413	.413	.423	.413	.413	.423	.452	.462	.519	
50	.404	.404	.394	.404	.394	.394	.442	.442	.490	
30	.558	.577	.519	.471	.423	.404	.452	.442	.490	
20	.683	.683	.615	.519	.423	.404	.423	.423	.442	
10	.769	.817	.788	.683	.538	.452	.423	.404	.452	
4	.625	.721	.760	.731	.644	.442	.423	.404	.452	
2	.615	.625	.702	.721	.663	.510	.433	.413	.452	
1	.519	.567	.654	.750	.654	.500	.423	.413	.462	

Table 15: valgN