

Table 4.5: Rank correlations of the inferred relevance judgements to the original document judgments reported in terms of Kendall's τ and Spearman's ρ .

Approach			Touché 20		Robust04		TREC-7		TREC-8		TREC-19 DL		TREC-20 DL	
			τ	ρ	τ	ρ	τ	ρ	τ	ρ	τ	ρ	τ	ρ
Pairwise Preferences flan-t5-base	mean	id	0.230	0.279	0.000	0.000	0.154	0.188	0.202	0.246	0.298	0.374	0.293	0.368
		log	0.230	0.279	0.000	0.000	0.154	0.188	0.202	0.246	0.298	0.374	0.293	0.368
		exp	0.230	0.279	0.000	0.000	0.154	0.188	0.202	0.246	0.298	0.374	0.293	0.368
		sqrt	0.230	0.279	0.000	0.000	0.154	0.188	0.202	0.246	0.298	0.374	0.293	0.368
	min	id	0.248	0.303	0.000	0.000	0.169	0.207	0.217	0.265	0.349	0.436	0.341	0.426
		log	0.248	0.303	0.000	0.000	0.169	0.207	0.217	0.265	0.349	0.436	0.341	0.426
		exp	0.248	0.303	0.000	0.000	0.169	0.207	0.217	0.265	0.349	0.436	0.341	0.426
		sqrt	0.248	0.303	0.000	0.000	0.169	0.207	0.217	0.265	0.349	0.436	0.341	0.426
	max	id	0.127	0.155	0.000	0.000	0.046	0.056	0.034	0.041	0.109	0.142	0.147	0.186
		log	0.127	0.155	0.000	0.000	0.046	0.056	0.034	0.041	0.109	0.142	0.147	0.186
		exp	0.127	0.155	0.000	0.000	0.046	0.056	0.034	0.041	0.109	0.142	0.147	0.186
		sqrt	0.127	0.155	0.000	0.000	0.046	0.056	0.034	0.041	0.109	0.142	0.147	0.186
	sum	id	0.230	0.279	0.000	0.000	0.154	0.188	0.202	0.246	0.298	0.374	0.293	0.368
		log	0.230	0.279	0.000	0.000	0.154	0.188	0.202	0.246	0.298	0.374	0.293	0.368
		exp	0.230	0.279	0.000	0.000	0.154	0.188	0.202	0.246	0.298	0.374	0.293	0.368
		sqrt	0.230	0.279	0.000	0.000	0.154	0.188	0.202	0.246	0.298	0.374	0.293	0.368
Pairwise Preferences flan-t5-small	mean	id	0.214	0.262	0.000	0.000	0.087	0.106	0.106	0.129	0.211	0.268	0.186	0.235
		log	0.214	0.262	0.000	0.000	0.087	0.106	0.106	0.129	0.211	0.268	0.186	0.235
		exp	0.214	0.262	0.000	0.000	0.087	0.106	0.106	0.129	0.211	0.268	0.186	0.235
		sqrt	0.214	0.262	0.000	0.000	0.087	0.106	0.106	0.129	0.211	0.268	0.186	0.235
	min	id	0.233	0.285	0.000	0.000	0.106	0.129	0.112	0.137	0.256	0.325	0.228	0.288
		log	0.233	0.285	0.000	0.000	0.106	0.129	0.112	0.137	0.256	0.325	0.228	0.288
		exp	0.233	0.285	0.000	0.000	0.106	0.129	0.112	0.137	0.256	0.325	0.228	0.288
		sqrt	0.233	0.285	0.000	0.000	0.106	0.129	0.112	0.137	0.256	0.325	0.228	0.288
	max	id	0.129	0.161	0.000	0.000	0.021	0.026	0.048	0.059	0.079	0.100	0.076	0.096
		log	0.129	0.161	0.000	0.000	0.021	0.026	0.048	0.059	0.079	0.100	0.076	0.096
		exp	0.129	0.161	0.000	0.000	0.021	0.026	0.048	0.059	0.079	0.100	0.076	0.096
		sqrt	0.129	0.161	0.000	0.000	0.021	0.026	0.048	0.059	0.079	0.100	0.076	0.096
	sum	id	0.214	0.262	0.000	0.000	0.087	0.106	0.106	0.129	0.211	0.268	0.186	0.235
		log	0.214	0.262	0.000	0.000	0.087	0.106	0.106	0.129	0.211	0.268	0.186	0.235
		exp	0.214	0.262	0.000	0.000	0.087	0.106	0.106	0.129	0.211	0.268	0.186	0.235
		sqrt	0.214	0.262	0.000	0.000	0.087	0.106	0.106	0.129	0.211	0.268	0.186	0.235
Pairwise Preferences t5-small	mean	id	0.113	0.140	0.000	0.000	-0.045	-0.055	-0.060	-0.073	-0.040	-0.052	-0.029	-0.038
		log	0.113	0.140	0.000	0.000	-0.045	-0.055	-0.060	-0.073	-0.040	-0.052	-0.029	-0.038
		exp	0.113	0.140	0.000	0.000	-0.045	-0.055	-0.060	-0.073	-0.040	-0.052	-0.029	-0.038
		sqrt	0.113	0.140	0.000	0.000	-0.045	-0.055	-0.060	-0.073	-0.040	-0.052	-0.029	-0.038
	min	id	0.151	0.186	0.000	0.000	-0.038	-0.046	-0.066	-0.080	-0.025	-0.034	-0.008	-0.011
		log	0.151	0.186	0.000	0.000	-0.038	-0.046	-0.066	-0.080	-0.025	-0.034	-0.008	-0.011
		exp	0.151	0.186	0.000	0.000	-0.038	-0.046	-0.066	-0.080	-0.025	-0.034	-0.008	-0.011
		sqrt	0.151	0.186	0.000	0.000	-0.038	-0.046	-0.066	-0.080	-0.025	-0.034	-0.008	-0.011
	max	id	0.041	0.048	0.000	0.000	-0.040	-0.048	-0.048	-0.059	-0.026	-0.034	-0.032	-0.041
		log	0.041	0.048	0.000	0.000	-0.040	-0.048	-0.048	-0.059	-0.026	-0.034	-0.032	-0.041
		exp	0.041	0.048	0.000	0.000	-0.040	-0.048	-0.048	-0.059	-0.026	-0.034	-0.032	-0.041
		sqrt	0.041	0.048	0.000	0.000	-0.040	-0.048	-0.048	-0.059	-0.026	-0.034	-0.032	-0.041
	sum	id	0.113	0.140	0.000	0.000	-0.045	-0.055	-0.060	-0.073	-0.040	-0.052	-0.029	-0.038
		log	0.113	0.140	0.000	0.000	-0.045	-0.055	-0.060	-0.073	-0.040	-0.052	-0.029	-0.038
		exp	0.113	0.140	0.000	0.000	-0.045	-0.055	-0.060	-0.073	-0.040	-0.052	-0.029	-0.038
		sqrt	0.113	0.140	0.000	0.000	-0.045	-0.055	-0.060	-0.073	-0.040	-0.052	-0.029	-0.038
Pointwise Preferences	flan-t5-base		0.149	0.179	0.000	0.000	-0.064	-0.078	-0.078	-0.095	0.026	0.034	0.000	0.001
	flan-t5-small		-0.044	-0.056	0.000	0.000	-0.048	-0.059	-0.072	-0.087	-0.037	-0.048	-0.015	-0.020
	t5-small		-0.068	-0.082	0.000	0.000	-0.047	-0.057	-0.065	-0.079	-0.040	-0.052	-0.021	-0.027

Table 4.6: Correlations of the inferred relevance judgements to the original document judgments reported in terms of Kendall’s τ and Spearman’s ρ .

Approach			Touché 20		TREC-7		TREC-8		TREC-19 DL		TREC-20 DL	
			τ	ρ	τ	ρ	τ	ρ	τ	ρ	τ	ρ
Pairwise Preferences flant5-base	mean	id	0.218	0.271	0.117	0.143	0.168	0.205	0.159	0.199	0.218	0.270
		log	0.218	0.271	0.117	0.143	0.168	0.205	0.159	0.199	0.218	0.270
		exp	0.218	0.271	0.117	0.143	0.168	0.205	0.159	0.199	0.218	0.270
		sqrt	0.218	0.271	0.117	0.143	0.168	0.205	0.159	0.199	0.218	0.270
	min	id	0.233	0.293	0.125	0.153	0.184	0.225	0.184	0.231	0.225	0.279
		log	0.233	0.293	0.125	0.153	0.184	0.225	0.184	0.231	0.225	0.279
		exp	0.233	0.293	0.125	0.153	0.184	0.225	0.184	0.231	0.225	0.279
		sqrt	0.233	0.293	0.125	0.153	0.184	0.225	0.184	0.231	0.225	0.279
	max	id	0.222	0.275	0.087	0.106	0.074	0.090	0.077	0.097	0.100	0.123
		log	0.222	0.275	0.087	0.106	0.074	0.090	0.077	0.097	0.100	0.123
		exp	0.222	0.275	0.087	0.106	0.074	0.090	0.077	0.097	0.100	0.123
		sqrt	0.222	0.275	0.087	0.106	0.074	0.090	0.077	0.097	0.100	0.123
	sum	id	0.218	0.271	0.117	0.143	0.168	0.205	0.159	0.199	0.218	0.270
		log	0.218	0.271	0.117	0.143	0.168	0.205	0.159	0.199	0.218	0.270
		exp	0.218	0.271	0.117	0.143	0.168	0.205	0.159	0.199	0.218	0.270
		sqrt	0.218	0.271	0.117	0.143	0.168	0.205	0.159	0.199	0.218	0.270
Pairwise Preferences flant5-small	mean	id	0.253	0.324	0.056	0.069	0.073	0.089	0.142	0.179	0.136	0.169
		log	0.253	0.324	0.056	0.069	0.073	0.089	0.142	0.179	0.136	0.169
		exp	0.253	0.324	0.056	0.069	0.073	0.089	0.142	0.179	0.136	0.169
		sqrt	0.253	0.324	0.056	0.069	0.073	0.089	0.142	0.179	0.136	0.169
	min	id	0.300	0.383	0.072	0.088	0.090	0.110	0.142	0.178	0.154	0.191
		log	0.300	0.383	0.072	0.088	0.090	0.110	0.142	0.178	0.154	0.191
		exp	0.300	0.383	0.072	0.088	0.090	0.110	0.142	0.178	0.154	0.191
		sqrt	0.300	0.383	0.072	0.088	0.090	0.110	0.142	0.178	0.154	0.191
	max	id	0.200	0.254	0.028	0.034	0.033	0.041	0.050	0.062	0.081	0.100
		log	0.200	0.254	0.028	0.034	0.033	0.041	0.050	0.062	0.081	0.100
		exp	0.200	0.254	0.028	0.034	0.033	0.041	0.050	0.062	0.081	0.100
		sqrt	0.200	0.254	0.028	0.034	0.033	0.041	0.050	0.062	0.081	0.100
	sum	id	0.253	0.324	0.056	0.069	0.073	0.089	0.142	0.179	0.136	0.169
		log	0.253	0.324	0.056	0.069	0.073	0.089	0.142	0.179	0.136	0.169
		exp	0.253	0.324	0.056	0.069	0.073	0.089	0.142	0.179	0.136	0.169
		sqrt	0.253	0.324	0.056	0.069	0.073	0.089	0.142	0.179	0.136	0.169
Pairwise Preferences t5-small	mean	id	0.250	0.322	0.000	0.000	-0.008	-0.010	0.003	0.005	-0.001	-0.001
		log	0.250	0.322	0.000	0.000	-0.008	-0.010	0.003	0.005	-0.001	-0.001
		exp	0.250	0.322	0.000	0.000	-0.008	-0.010	0.003	0.005	-0.001	-0.001
		sqrt	0.250	0.322	0.000	0.000	-0.008	-0.010	0.003	0.005	-0.001	-0.001
	min	id	0.310	0.400	0.017	0.021	0.005	0.007	0.007	0.010	0.013	0.016
		log	0.310	0.400	0.017	0.021	0.005	0.007	0.007	0.010	0.013	0.016
		exp	0.310	0.400	0.017	0.021	0.005	0.007	0.007	0.010	0.013	0.016
		sqrt	0.310	0.400	0.017	0.021	0.005	0.007	0.007	0.010	0.013	0.016
	max	id	0.164	0.213	0.004	0.005	-0.005	-0.007	0.005	0.007	0.008	0.010
		log	0.164	0.213	0.004	0.005	-0.005	-0.007	0.005	0.007	0.008	0.010
		exp	0.164	0.213	0.004	0.005	-0.005	-0.007	0.005	0.007	0.008	0.010
		sqrt	0.164	0.213	0.004	0.005	-0.005	-0.007	0.005	0.007	0.008	0.010
	sum	id	0.250	0.322	0.000	0.000	-0.008	-0.010	0.003	0.005	-0.001	-0.001
		log	0.250	0.322	0.000	0.000	-0.008	-0.010	0.003	0.005	-0.001	-0.001
		exp	0.250	0.322	0.000	0.000	-0.008	-0.010	0.003	0.005	-0.001	-0.001
		sqrt	0.250	0.322	0.000	0.000	-0.008	-0.010	0.003	0.005	-0.001	-0.001
Pointwise Preferences	flant5-base		0.222	0.275	0.087	0.106	0.074	0.090	0.077	0.097	0.100	0.123
	flant5-small		0.200	0.254	0.028	0.034	0.033	0.041	0.050	0.062	0.081	0.100
	t5-small		0.164	0.213	0.004	0.005	-0.005	-0.007	0.005	0.007	0.008	0.010

APPENDIX A. GREEDY EVALUATION

Table A.2: Rank correlations of the inferred relevance judgements to the original document judgments reported in terms of the greedy version of Kendall’s τ and Spearman’s ρ .

Approach		Touché 20		Robust04		TREC-7		TREC-8		TREC-19 DL		TREC-20 DL		
		τ	ρ	τ	ρ	τ	ρ	τ	ρ	τ	ρ	τ	ρ	
Pairwise Preferences flan-t5-base	mean	id	0.477	0.487	0.000	0.000	0.306	0.306	0.343	0.343	0.403	0.430	0.443	0.467
		log	0.475	0.485	0.000	0.000	0.304	0.304	0.340	0.340	0.404	0.431	0.440	0.463
		exp	0.474	0.484	0.000	0.000	0.305	0.305	0.340	0.340	0.400	0.427	0.444	0.468
		sqrt	0.474	0.484	0.000	0.000	0.305	0.305	0.341	0.341	0.399	0.426	0.441	0.465
	min	id	0.468	0.481	0.000	0.000	0.331	0.331	0.360	0.360	0.456	0.486	0.483	0.510
		log	0.467	0.479	0.000	0.000	0.330	0.330	0.358	0.358	0.456	0.486	0.484	0.510
		exp	0.467	0.479	0.000	0.000	0.332	0.332	0.360	0.360	0.456	0.485	0.487	0.513
		sqrt	0.468	0.481	0.000	0.000	0.329	0.329	0.359	0.359	0.456	0.485	0.484	0.510
	max	id	0.377	0.384	0.000	0.000	0.171	0.171	0.178	0.178	0.226	0.240	0.314	0.329
		log	0.372	0.378	0.000	0.000	0.174	0.174	0.179	0.179	0.228	0.242	0.315	0.330
		exp	0.374	0.381	0.000	0.000	0.171	0.171	0.179	0.179	0.227	0.241	0.315	0.330
		sqrt	0.373	0.379	0.000	0.000	0.174	0.174	0.177	0.177	0.227	0.241	0.309	0.324
	sum	id	0.477	0.487	0.000	0.000	0.306	0.306	0.343	0.343	0.403	0.430	0.443	0.467
		log	0.475	0.485	0.000	0.000	0.304	0.304	0.340	0.340	0.404	0.431	0.440	0.463
		exp	0.389	0.398	0.000	0.000	0.284	0.284	0.318	0.318	0.353	0.377	0.412	0.434
		sqrt	0.474	0.484	0.000	0.000	0.305	0.305	0.341	0.341	0.399	0.426	0.441	0.465
Pairwise Preferences flan-t5-small	mean	id	0.413	0.422	0.000	0.000	0.236	0.236	0.253	0.253	0.304	0.325	0.312	0.329
		log	0.410	0.418	0.000	0.000	0.236	0.236	0.252	0.252	0.304	0.324	0.312	0.330
		exp	0.413	0.422	0.000	0.000	0.234	0.234	0.250	0.250	0.301	0.320	0.313	0.331
		sqrt	0.411	0.421	0.000	0.000	0.237	0.237	0.251	0.251	0.303	0.323	0.312	0.330
	min	id	0.436	0.446	0.000	0.000	0.253	0.253	0.266	0.266	0.360	0.384	0.354	0.373
		log	0.440	0.451	0.000	0.000	0.255	0.255	0.267	0.267	0.355	0.379	0.355	0.375
		exp	0.439	0.449	0.000	0.000	0.254	0.254	0.267	0.267	0.356	0.379	0.354	0.373
		sqrt	0.442	0.452	0.000	0.000	0.254	0.254	0.265	0.265	0.359	0.382	0.354	0.373
	max	id	0.331	0.339	0.000	0.000	0.172	0.172	0.193	0.193	0.169	0.181	0.201	0.212
		log	0.330	0.338	0.000	0.000	0.173	0.173	0.193	0.193	0.169	0.181	0.194	0.204
		exp	0.330	0.338	0.000	0.000	0.172	0.172	0.193	0.193	0.170	0.182	0.193	0.204
		sqrt	0.333	0.341	0.000	0.000	0.174	0.174	0.193	0.193	0.164	0.175	0.193	0.203
	sum	id	0.413	0.422	0.000	0.000	0.236	0.236	0.253	0.253	0.304	0.325	0.312	0.329
		log	0.410	0.418	0.000	0.000	0.236	0.236	0.252	0.252	0.304	0.324	0.312	0.330
		exp	0.410	0.419	0.000	0.000	0.232	0.232	0.248	0.248	0.305	0.326	0.312	0.330
		sqrt	0.411	0.421	0.000	0.000	0.237	0.237	0.251	0.251	0.303	0.323	0.312	0.330
Pairwise Preferences t5-small	mean	id	0.354	0.358	0.000	0.000	0.117	0.117	0.116	0.116	0.051	0.053	0.045	0.045
		log	0.351	0.356	0.000	0.000	0.116	0.116	0.115	0.115	0.048	0.050	0.043	0.042
		exp	0.352	0.357	0.000	0.000	0.116	0.116	0.116	0.116	0.053	0.055	0.047	0.047
		sqrt	0.354	0.360	0.000	0.000	0.116	0.116	0.116	0.116	0.048	0.051	0.039	0.039
	min	id	0.392	0.400	0.000	0.000	0.119	0.119	0.097	0.097	0.065	0.069	0.085	0.088
		log	0.387	0.395	0.000	0.000	0.121	0.121	0.100	0.100	0.061	0.065	0.084	0.087
		exp	0.389	0.396	0.000	0.000	0.119	0.119	0.098	0.098	0.059	0.062	0.082	0.084
		sqrt	0.389	0.397	0.000	0.000	0.120	0.120	0.099	0.099	0.063	0.066	0.081	0.084
	max	id	0.276	0.281	0.000	0.000	0.130	0.130	0.116	0.116	0.062	0.065	0.091	0.095
		log	0.280	0.285	0.000	0.000	0.129	0.129	0.115	0.115	0.065	0.069	0.088	0.091
		exp	0.279	0.283	0.000	0.000	0.131	0.131	0.117	0.117	0.057	0.060	0.086	0.089
		sqrt	0.276	0.280	0.000	0.000	0.131	0.131	0.117	0.117	0.060	0.063	0.090	0.093
	sum	id	0.354	0.358	0.000	0.000	0.117	0.117	0.116	0.116	0.051	0.053	0.045	0.045
		log	0.351	0.356	0.000	0.000	0.116	0.116	0.115	0.115	0.048	0.050	0.043	0.042
		exp	0.323	0.329	0.000	0.000	0.094	0.094	0.094	0.094	0.050	0.053	0.064	0.067
		sqrt	0.354	0.360	0.000	0.000	0.116	0.116	0.116	0.116	0.048	0.051	0.039	0.039
Pointwise Preferences	flan-t5-base	0.363	0.370	0.000	0.000	0.094	0.094	0.106	0.106	0.111	0.116	0.071	0.074	
	flan-t5-small	0.193	0.196	0.000	0.000	0.104	0.104	0.100	0.100	0.038	0.039	0.092	0.096	
	t5-small	0.201	0.203	0.000	0.000	0.098	0.098	0.107	0.107	0.044	0.046	0.068	0.072	