

TABLE II. SIMILARITY RESULTS FOR WHOLE TEXT OF STANDARD DOCUMENTS BY DOMAIN EXPERTS

File Name	Jaccard Similarity				Cosine Similarity				Cosine with Word Vector Similarity			
	Yake	TopicRank	MultipartiteRank	KPMiner	Yake	TopicRank	MultipartiteRank	KPMiner	Yake	TopicRank	MultipartiteRank	KPMiner
File-1	0.081	0.097	0.153	0.097	0.170	0.178	0.269	0.190	0.868	0.904	0.912	0.891
File-2	0.074	0.097	0.114	0.079	0.163	0.178	0.207	0.155	0.831	0.901	0.900	0.886
File-3	0.065	0.080	0.070	0.097	0.139	0.152	0.133	0.190	0.828	0.924	0.910	0.858
File-4	0.097	0.116	0.133	0.108	0.182	0.210	0.238	0.206	0.862	0.912	0.922	0.917
File-5	0.139	0.104	0.116	0.117	0.272	0.190	0.210	0.217	0.878	0.884	0.882	0.916
File-6	0.095	0.160	0.178	0.134	0.190	0.278	0.305	0.260	0.904	0.937	0.942	0.909
File-7	0.062	0.169	0.143	0.110	0.129	0.293	0.254	0.210	0.826	0.920	0.930	0.876
File-8	0.063	0.065	0.082	0.082	0.131	0.124	0.155	0.163	0.863	0.914	0.886	0.883
File-9	0.155	0.169	0.136	0.107	0.294	0.293	0.244	0.203	0.923	0.924	0.895	0.941
File-10	0.147	0.116	0.159	0.145	0.294	0.211	0.278	0.272	0.895	0.933	0.927	0.898

TABLE III. SIMILARITY RESULTS FOR POSITIVE TEXT OF STANDARD DOCUMENTS BY DOMAIN EXPERTS

File Name	Jaccard Similarity				Cosine Similarity				Cosine with Word Vector Similarity			
	Yake	TopicRank	MultipartiteRank	KPMiner	Yake	TopicRank	MultipartiteRank	KPMiner	Yake	TopicRank	MultipartiteRank	KPMiner
File-1	0.072	0.135	0.169	0.078	0.145	0.240	0.293	0.152	0.814	0.898	0.926	0.913
File-2	0.081	0.094	0.086	0.095	0.172	0.173	0.160	0.183	0.809	0.921	0.897	0.884
File-3	0.065	0.056	0.095	0.136	0.132	0.107	0.178	0.266	0.809	0.924	0.899	0.890
File-4	0.084	0.156	0.133	0.129	0.147	0.272	0.238	0.245	0.785	0.924	0.926	0.924
File-5	0.086	0.126	0.122	0.137	0.169	0.226	0.220	0.253	0.828	0.890	0.857	0.908
File-6	0.084	0.156	0.174	0.087	0.173	0.272	0.300	0.177	0.878	0.926	0.921	0.877
File-7	0.067	0.149	0.143	0.141	0.137	0.263	0.254	0.262	0.780	0.920	0.918	0.912
File-8	0.082	0.065	0.045	0.110	0.166	0.124	0.089	0.210	0.871	0.902	0.883	0.839
File-9	0.143	0.155	0.149	0.111	0.274	0.272	0.263	0.214	0.884	0.930	0.937	0.923
File-10	0.157	0.160	0.185	0.120	0.318	0.282	0.317	0.224	0.853	0.929	0.930	0.824

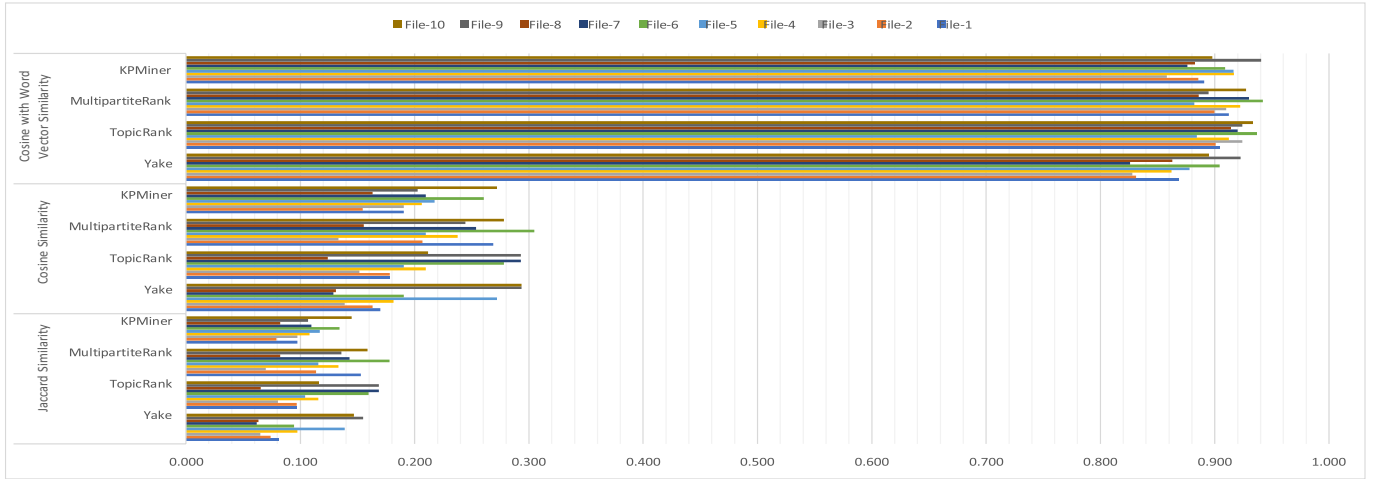


Fig. 2. Score comparison for different similarity calculation algorithms for whole text of documents

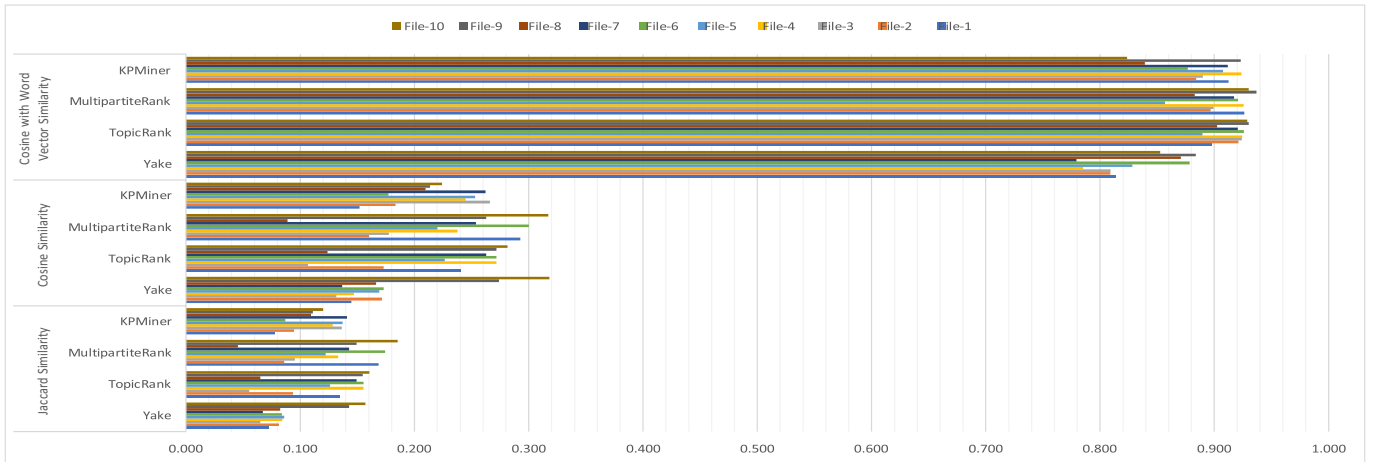


Fig. 3. Score comparison for different similarity calculation algorithms for positive text of documents