

Out-Expansion Complete Results

Type	Acronym Out-Expansion Technique		Out-Expansion Accuracy (%)				Execution time per document (s)			
	Predictor	Representator	MSH	Science WISE	CS Wiki	AVG	MSH	Science WISE	CS Wiki	AVG
Related Work	CosSim	Classic Context Vector	97.93%	93.98%	80.47%	90.79%	0.1087	0.0219	0.1363	0.0890
	CosSim	SBE Li et al.	94.88%	91.55%	66.44%	84.29%	0.1182	0.0203	0.1266	0.0884
	CosSim	Thakker et al.	93.49%	91.29%	74.42%	86.40%	7.5484	2.1189	3.2777	4.3150
	CosSim	UAD SBE	90.51%	60.84%	42.37%	64.57%	0.0530	0.0053	0.1086	0.0556
	CosSim	UAD Vote	89.18%	61.26%	43.33%	64.59%	0.0529	0.0052	0.1100	0.0560
Direct Representator	CosSim	TFIDF	80.08%	70.64%	81.02%	77.25%	0.0721	0.0112	0.1286	0.0706
	CosSim	LDA	90.87%	89.99%	61.21%	80.69%	0.0527	0.0053	0.1129	0.0570
	CosSim	Doc2Vec	97.25%	93.98%	79.23%	90.15%	0.0678	0.0102	0.1921	0.0900
	SVM	TFIDF	83.91%	80.76%	82.53%	82.40%	0.0737	0.0111	0.1334	0.0727
	LR	TFIDF	84.20%	80.08%	82.81%	82.36%	0.0730	0.0111	0.1367	0.0736
	RF	TFIDF	84.42%	79.98%	62.86%	75.75%	0.7616	0.0257	2.4412	1.0762
	SVM	LDA	91.84%	94.24%	70.84%	85.64%	0.0533	0.0057	0.1126	0.0572
	LR	LDA	91.65%	94.04%	70.29%	85.32%	0.0533	0.0057	0.1136	0.0575
	RF	LDA	91.39%	89.52%	67.13%	82.68%	0.2292	0.0444	0.6456	0.3064
	SVM	Doc2Vec	97.60%	94.24%	81.02%	90.95%	0.0683	0.0104	0.1935	0.0907
	LR	Doc2Vec	97.80%	93.67%	80.61%	90.69%	0.0685	0.0106	0.1966	0.0919
	RF	Doc2Vec	95.01%	85.22%	68.09%	82.77%	0.6229	0.3958	0.7388	0.5858
	CosSim	Document Context Vector	96.70%	92.27%	80.19%	89.72%	0.3068	0.0587	0.2757	0.2137
	Document Context Vector	SVM	Document Context Vector	97.09%	81.38%	66.44%	81.63%	0.3009	0.0582	0.3237
LR		Document Context Vector	94.95%	81.64%	66.16%	80.92%	0.3245	0.0697	0.5177	0.3039
RF		Document Context Vector	95.50%	80.65%	55.57%	77.24%	4.8147	1.6559	25.9610	10.8106
Combination of Representators	CosSim	Context Vector + Doc2Vec	97.28%	94.04%	79.23%	90.18%	0.1755	0.0405	0.2315	0.1492
	SVM	Context Vector + Doc2Vec	97.60%	94.40%	81.02%	91.01%	0.1925	0.0450	0.2827	0.1734
	LR	Context Vector + Doc2Vec	97.60%	92.48%	80.61%	90.23%	0.2630	0.0633	0.6816	0.3360
	RF	Context Vector + Doc2Vec	40.90%	50.41%	36.04%	42.45%	0.1836	0.0471	25.0791	8.4366
	CosSim	Document Context Vector + Doc2Vec	97.25%	94.09%	79.23%	90.19%	0.3686	0.0751	0.3762	0.2733
	SVM	Document Context Vector + Doc2Vec	97.67%	94.24%	81.02%	90.98%	0.3741	0.0766	0.4211	0.2906
	LR	Document Context Vector + Doc2Vec	97.80%	93.57%	80.61%	90.66%	0.4009	0.0851	0.8006	0.4289
RF	Document Context Vector + Doc2Vec	93.98%	80.96%	60.80%	78.58%	3.7988	0.9837	24.9490	9.9105	

Models Execution Times

Models	Models Training Execution Times			
	MSH	Science WISE	CS Wiki	AVG
Classic Context Vector	1.23	0.34	5.11	2.22
Word2Vec used by SBE (Li et al.)	23.58	7.28	156.68	62.51
Word2Vec used by UADs SBE/Vote	96.39	54.35	473.48	208.07
TFIDF	6.11	1.42	6.40	4.64
LDA	491.91	201.36	1413.10	702.12
Doc2Vec	110.95	19.49	148.31	92.92
Document Context Vector	8.91	2.82	43.27	18.33

Out-Expansion P-Values

P-values of Statistical Paired Test for all vs SVM with Doc2Vec + Classic Context Vector								
Type	Acronym Out-Expansion Technique		Out-Expansion Accuracy P-value			Execution time per document P-value		
	Predictor	Representator	MSH	Science WISE	CS Wiki	MSH	Science WISE	CS Wiki
Related Work	Cossim	Classic Context Vector	0.199	0.255	0.374	0	0	0
	Cossim	SBE Li et al.	0	0	0	0	0	0
	Cossim	Thakker et al.	0	0	0	0	0	0
	Cossim	UAD SBE	0	0	0	0	0	0
	Cossim	UAD Vote	0	0	0	0	0	0
Direct Representator	Cossim	TFIDF	0	0	0.548	0	0	0
	Cossim	LDA	0	0	0	0	0	0
	Cossim	Doc2Vec	0.118	0.142	0.04	0	0	0
Machine Learning as Expansion Predictor	SVM	TFIDF	0	0	0.089	0	0	0
	LR	TFIDF	0	0	0.054	0	0	0
	RF	TFIDF	0	0	0	0	0	0
	SVM	LDA	0	0.434	0	0	0	0
	LR	LDA	0	0.306	0	0	0	0
	RF	LDA	0	0	0	0	0.259	0
	SVM	Doc2Vec	0.584	0.218	1	0	0	0
	LR	Doc2Vec	0.113	0.002	0.228	0	0	0
Document Context Vector	RF	Doc2Vec	0	0	0	0	0	0
	Cossim	Document Context Vector	0.012	0	0.291	0	0	0
	SVM	Document Context Vector	0.071	0	0	0	0	0
	LR	Document Context Vector	0	0	0	0	0	0
Combination of Representators	RF	Document Context Vector	0	0	0	0	0	0
	Cossim	Context Vector + Doc2Vec	0.143	0.16	0.04	0	0	0
	LR	Context Vector + Doc2Vec	0.583	0	0.228	0	0	0
	RF	Context Vector + Doc2Vec	0	0	0	0	0	0
	Cossim	Document Context Vector + Doc2Vec	0.118	0.219	0.04	0	0	0.001
	SVM	Document Context Vector + Doc2Vec	0.422	0.218	1	0	0	0
	LR	Document Context Vector + Doc2Vec	0.113	0	0.228	0	0	0
RF	Document Context Vector + Doc2Vec	0	0	0	0	0	0	

End-to-end P-values

P-values of Statistical Paired Test Matrix for F1-Measure			
	Predictor	SVM	SVM
Predictor	Representator/Representator	Doc2Vec	Context Vector + Doc2Vec
CosSim	Classic Context Vector		0
SVM	Doc2Vec	1	0.512