

Supplementary Material to “Analytical Methods for Interpretable Ultradense Word Embeddings”

Philipp Dufter, Hinrich Schütze

Center for Information and Language Processing (CIS)

LMU Munich, Germany

philipp@cis.lmu.de

1 Code

The code which was used to conduct the experiments in this paper is available at <https://github.com/pdufter/densray>.

2 Continuous Lexicon

In case of a continuous lexicon $l : V \rightarrow \mathbb{R}$ one can extend Equation 2 in the main paper by defining:

$$A := \sum_{(v,w) \in V \times V} -l(v)l(w)d_{vw}d_{vw}^T$$

In the case of a binary lexicon Equation 2 from the main paper is recovered for $\alpha_{\neq} = \alpha_{=} = 1$.

3 Full Analogy Results

In this section we present the results of the word analogy task per category. See Table 1 and Table 2 for detailed results with the methods IntCos Complement and Original, respectively. The format and numbers presented are the same as in the corresponding table from the main paper.

		FastText					Google News					
Google Analogy		Mean Cosine Sim			Precision		Mean Cosine Sim			Precision		
		Inter	Intra	R	IntCos	LRCos	Inter	Intra	R	IntCos	LRCos	
		IntraL	IntraR		DensRaySVM		IntraL	IntraR		DensRaySVM		
	capital-common-countries	0.76	0.53	0.56	1.00	1.00	1.00			0.91	0.96	0.91
	capital-world	0.75	0.44	0.51	0.97	0.96	1.00			0.86	0.88	0.90
	city-in-state	0.71	0.51	0.63	0.78	0.79	0.85			0.82	0.85	0.87
	currency	0.33	0.59	0.48	0.62	0.69	0.08			0.37	0.42	0.44
	family	0.84	0.57	0.59	0.91	0.91	0.95			0.74	0.48	0.53
	gram1-adjective-to-adverb	0.66	0.50	0.56	0.78	0.84	0.88			0.49	0.21	0.26
	gram2-opposite	0.72	0.51	0.54	0.69	0.83	0.76			0.50	0.24	0.31
	gram3-comparative	0.75	0.53	0.57	0.84	0.89	0.95			0.60	0.25	0.40
	gram4-superlative	0.70	0.53	0.61	0.94	1.00	1.00			0.54	0.26	0.39
	gram5-present-participle	0.76	0.43	0.48	1.00	0.94	1.00			0.70	0.20	0.21
	gram6-nationality-adjective	0.70	0.55	0.54	0.90	0.90	0.93			0.72	0.41	0.41
	gram7-past-tense	0.73	0.49	0.47	0.82	0.90	0.97			0.66	0.21	0.22
	gram8-plural	0.80	0.41	0.42	1.00	1.00	1.00			0.73	0.21	0.23
	gram9-plural-verbs	0.74	0.43	0.48	0.90	0.97	1.00			0.64	0.22	0.24
	Micro Mean	0.73	0.48	0.53	0.88	0.91	0.93			0.62	0.31	0.36
	Macro Mean	0.71	0.50	0.53	0.87	0.90	0.89			0.61	0.30	0.35
	Macro Std	0.11	0.05	0.06	0.11	0.08	0.23			0.10	0.09	0.10
	Derivational	0.63	0.47	0.45	0.74	0.78	0.80			0.44	0.21	0.20
	D01 [noun+less-reg]	0.51	0.36	0.45	0.50	0.53	0.60			0.26	0.16	0.24
	D02 [un+adj-reg]	0.71	0.46	0.46	0.68	0.72	0.84			0.47	0.17	0.20
	D03 [adj+ly-reg]	0.63	0.50	0.52	0.86	0.88	0.90			0.48	0.17	0.22
	D04 [over+adj-reg]	0.63	0.45	0.45	0.59	0.69	0.62			0.39	0.17	0.21
	D05 [adj+ness-reg]	0.63	0.49	0.51	0.92	1.00	0.92			0.47	0.21	0.26
	D06 [re+verb-reg]	0.74	0.52	0.49	0.50	0.62	0.76			0.56	0.29	0.28
	D07 [verb+able-reg]	0.57	0.48	0.45	0.71	0.66	0.63			0.38	0.25	0.20
	D08 [verb+er-irreg]	0.55	0.48	0.41	0.84	0.88	0.79			0.30	0.24	0.17
	D09 [verb+tion-irreg]	0.63	0.46	0.41	0.86	0.88	0.86			0.51	0.22	0.15
	D10 [verb+ment-irreg]	0.64	0.46	0.42	0.86	0.88	0.90			0.47	0.24	0.15
	Encyclopedia	0.48	0.43	0.55	0.30	0.43	0.45			0.35	0.29	0.42
	E01 [country - capital]	0.63	0.47	0.41	0.72	0.96	0.98			0.61	0.35	0.32
	E02 [country - language]	0.40	0.43	0.59	0.24	0.35	0.33			0.36	0.31	0.45
	E03 [UK-city - county]	0.59	0.49	0.57	0.22	0.36	0.36			0.41	0.36	0.52
	E04 [name - nationality]	0.28	0.39	0.64	0.45	0.66	0.60			0.20	0.20	0.39
	E05 [name - occupation]	0.44	0.41	0.57	0.42	0.65	0.73			0.33	0.21	0.40
	E06 [animal - young]	0.47	0.43	0.44	0.05	0.07	0.15			0.34	0.36	0.38
	E07 [animal - sound]	0.37	0.43	0.40	0.15	0.20	0.22			0.15	0.31	0.25
	E08 [animal - shelter]	0.44	0.42	0.51	0.00	0.07	0.13			0.25	0.29	0.39
	E09 [things - color]	0.44	0.38	0.81	0.04	0.22	0.16			0.20	0.23	0.63
	E10 [male - female]	0.73	0.43	0.43	0.68	0.68	0.78			0.62	0.28	0.33
	Inflectional	0.75	0.48	0.51	0.92	0.93	0.97			0.63	0.22	0.23
	I01 [noun - plural-reg]	0.79	0.39	0.41	0.98	1.00	1.00			0.69	0.13	0.16
	I02 [noun - plural-irreg]	0.77	0.40	0.42	0.80	0.80	0.84			0.62	0.12	0.16
	I03 [adj - comparative]	0.75	0.50	0.52	0.97	1.00	1.00			0.63	0.23	0.37
	I04 [adj - superlative]	0.71	0.51	0.58	0.96	0.96	1.00			0.59	0.26	0.39
	I05 [verb-inf - 3pSg]	0.77	0.52	0.53	0.96	0.98	1.00			0.65	0.26	0.33
	I06 [verb-inf - Ving]	0.77	0.51	0.51	0.80	0.88	0.96			0.67	0.26	0.19
	I07 [verb-inf - Ved]	0.75	0.51	0.55	0.92	0.94	1.00			0.66	0.25	0.20
	I08 [verb-Ving - 3pSg]	0.70	0.48	0.51	0.94	0.96	0.98			0.56	0.17	0.31
	I09 [verb-Ving - Ved]	0.73	0.50	0.54	0.92	0.88	0.98			0.64	0.18	0.19
	I10 [verb-3pSg - Ved]	0.72	0.53	0.55	0.96	0.96	1.00			0.62	0.33	0.20
	Lexicography	0.62	0.37	0.38	0.17	0.20	0.26			0.45	0.17	0.18
	L01 [hypernyms - animals]	0.58	0.43	0.53	0.02	0.29	0.24			0.47	0.32	0.47
	L02 [hypernyms - misc]	0.55	0.35	0.39	0.14	0.11	0.14			0.42	0.21	0.21
	L03 [hyponyms - misc]	0.63	0.35	0.31	0.28	0.28	0.28			0.52	0.15	0.15
	L04 [meronyms - substance]	0.53	0.36	0.44	0.15	0.21	0.17			0.35	0.16	0.24
	L05 [meronyms - member]	0.58	0.36	0.36	0.10	0.10	0.12			0.36	0.15	0.15
	L06 [meronyms - part]	0.53	0.31	0.30	0.04	0.09	0.09			0.34	0.14	0.12
	L07 [synonyms - intensity]	0.67	0.37	0.37	0.25	0.25	0.36			0.51	0.17	0.16
	L08 [synonyms - exact]	0.71	0.33	0.31	0.18	0.16	0.22			0.55	0.11	0.11
	L09 [antonyms - gradable]	0.68	0.45	0.43	0.35	0.33	0.55			0.45	0.18	0.19
	L10 [antonyms - binary]	0.72	0.40	0.40	0.18	0.18	0.39			0.50	0.14	0.15
	Micro Mean	0.62	0.44	0.47	0.52	0.58	0.61			0.47	0.22	0.26
	Macro Mean	0.62	0.44	0.47	0.53	0.58	0.61			0.46	0.22	0.26
	Macro Std	0.12	0.06	0.09	0.34	0.33	0.32			0.14	0.07	0.12
	Derivational	0.44	0.21	0.20	0.55	0.50	0.44			0.26	0.16	0.24
	D01 [noun+less-reg]	0.47	0.17	0.20	0.66	0.54	0.58			0.48	0.17	0.22
	D02 [un+adj-reg]	0.48	0.17	0.22	0.70	0.76	0.76			0.41	0.44	0.30
	D03 [adj+ly-reg]	0.39	0.17	0.21	0.41	0.44	0.30			0.47	0.21	0.26
	D04 [over+adj-reg]	0.56	0.29	0.28	0.64	0.53	0.58			0.56	0.29	0.28
	D05 [adj+ness-reg]	0.38	0.25	0.20	0.38	0.28	0.19			0.38	0.25	0.20
	D06 [re+verb-reg]	0.30	0.24	0.17	0.29	0.19	0.07			0.30	0.24	0.17
	D07 [verb+able-reg]	0.51	0.22	0.15	0.73	0.63	0.51			0.51	0.22	0.15
	D08 [verb+er-irreg]	0.47	0.24	0.15	0.60	0.56	0.44			0.47	0.24	0.15
	D09 [verb+tion-irreg]	0.35	0.29	0.42	0.33	0.35	0.34			0.35	0.29	0.42
	D10 [verb+ment-irreg]	0.61	0.35	0.32	0.88	0.90	0.90			0.61	0.35	0.32
	Encyclopedia	0.36	0.31	0.45	0.47	0.30	0.36			0.36	0.31	0.45
	E01 [country - capital]	0.41	0.36	0.52	0.14	0.14	0.14			0.41	0.36	0.52
	E02 [country - language]	0.20	0.20	0.39	0.33	0.33	0.26			0.20	0.20	0.39
	E03 [UK-city - county]	0.33	0.21	0.40	0.45	0.62	0.52			0.33	0.21	0.40
	E04 [name - nationality]	0.34	0.36	0.38	0.06	0.06	0.12			0.34	0.36	0.38
	E05 [name - occupation]	0.15	0.31	0.25	0.17	0.03	0.00			0.15	0.31	0.25
	E06 [animal - young]	0.25	0.29	0.39	0.00	0.16	0.09			0.25	0.29	0.39
	E07 [animal - sound]	0.20	0.23	0.63	0.08	0.15	0.21			0.20	0.23	0.63
	E08 [animal - shelter]	0.62	0.28	0.33	0.66	0.68	0.68			0.62	0.28	0.33
	E09 [things - color]	0.63	0.22	0.23	0.88	0.87	0.88			0.63	0.22	0.23
	E10 [male - female]	0.69	0.13	0.16	0.84	0.84	0.88			0.69	0.13	0.16
	Inflectional	0.62	0.12	0.16	0.67	0.69	0.75			0.62	0.12	0.16
	I01 [noun - plural-reg]	0.63	0.23	0.37	0.97	0.97	1.00			0.63	0.23	0.37
	I02 [noun - plural-irreg]	0.59	0.26	0.39	0.93	0.93	0.97			0.59	0.26	0.39
	I03 [adj - comparative]	0.65	0.26	0.33	1.00	1.00	1.00			0.65	0.26	0.33
	I04 [adj - superlative]	0.67	0.26	0.19	0.84	0.82	0.82			0.67	0.26	0.19
	I05 [verb-inf - 3pSg]	0.66	0.25	0.20	0.92	0.90	0.88			0.66	0.25	0.20
	I06 [verb-inf - Ving]	0.56	0.17	0.31	0.90	0.92	0.90			0.56	0.17	0.31
	I07 [verb-inf - Ved]	0.64	0.18	0.19	0.84	0.84	0.82			0.64	0.18	0.19</

		FastText					Google News							
Google Analogy		Mean Cosine Sim			Precision		LRCos		Mean Cosine Sim			Precision		LRCos
		InterIntraL	IntraR	IntCos	DensRaySVM	IntCos		InterIntraL	IntraR	IntCos	DensRaySVM			
	capital-common-countries	0.76	0.53	0.56	1.00	1.00	1.00		0.64	0.37	0.38	0.91	0.91	0.91
	capital-world	0.75	0.44	0.51	0.98	0.98	1.00		0.64	0.34	0.36	0.90	0.90	0.90
	city-in-state	0.71	0.51	0.63	0.87	0.85	0.85		0.59	0.37	0.49	0.85	0.88	0.87
	currency	0.33	0.59	0.48	0.54	0.31	0.08		0.37	0.42	0.44	0.67	0.50	0.56
	family	0.84	0.57	0.59	0.91	0.91	0.95		0.74	0.48	0.53	0.87	0.87	0.87
	gram1-adjective-to-adverb	0.66	0.50	0.56	0.88	0.84	0.88		0.49	0.21	0.26	0.78	0.75	0.75
	gram2-opposite	0.72	0.51	0.54	0.72	0.83	0.76		0.50	0.24	0.31	0.71	0.75	0.71
	gram3-comparative	0.75	0.53	0.57	0.92	0.95	0.95		0.60	0.25	0.40	0.97	0.95	0.97
	gram4-superlative	0.70	0.53	0.61	1.00	1.00	1.00		0.54	0.26	0.39	0.94	0.94	0.94
	gram5-present-participle	0.76	0.43	0.48	1.00	1.00	1.00		0.70	0.20	0.21	0.94	0.91	0.94
	gram6-nationality-adjective	0.70	0.55	0.54	0.93	0.93	0.93		0.72	0.41	0.41	0.93	0.93	0.93
	gram7-past-tense	0.73	0.49	0.47	0.95	0.93	0.97		0.66	0.21	0.22	0.88	0.85	0.85
	gram8-plural	0.80	0.41	0.42	1.00	1.00	1.00		0.73	0.21	0.23	0.97	0.89	0.97
	gram9-plural-verbs	0.74	0.43	0.48	1.00	1.00	1.00		0.64	0.22	0.24	0.93	0.87	0.93
	Micro Mean	0.73	0.48	0.53	0.93	0.92	0.93		0.62	0.31	0.36	0.89	0.87	0.88
	Macro Mean	0.71	0.50	0.53	0.91	0.90	0.89		0.61	0.30	0.35	0.88	0.85	0.87
	Macro Std	0.11	0.05	0.06	0.12	0.17	0.23		0.10	0.09	0.10	0.09	0.11	0.11

		FastText					Google News							
BATS		Mean Cosine Sim			Precision		LRCos		Mean Cosine Sim			Precision		LRCos
		InterIntraL	IntraR	IntCos	DensRaySVM	IntCos		InterIntraL	IntraR	IntCos	DensRaySVM			
	Derivational	0.63	0.47	0.45	0.81	0.80	0.80		0.44	0.21	0.20	0.51	0.48	0.44
	D01 [noun+less-reg]	0.51	0.36	0.45	0.60	0.60	0.60		0.26	0.16	0.24	0.10	0.10	0.05
	D02 [un+adj-reg]	0.71	0.46	0.46	0.80	0.72	0.84		0.47	0.17	0.20	0.66	0.58	0.58
	D03 [adj+ly-reg]	0.63	0.50	0.52	0.90	0.90	0.90		0.48	0.17	0.22	0.76	0.78	0.76
	D04 [over+adj-reg]	0.63	0.45	0.45	0.66	0.69	0.62		0.39	0.17	0.21	0.41	0.41	0.30
	D05 [adj+ness-reg]	0.63	0.49	0.51	1.00	0.96	0.92		0.47	0.21	0.26	0.75	0.70	0.65
	D06 [re+verb-reg]	0.74	0.52	0.49	0.59	0.71	0.76		0.56	0.29	0.28	0.69	0.61	0.58
	D07 [verb+able-reg]	0.57	0.48	0.45	0.74	0.63	0.63		0.38	0.25	0.20	0.28	0.22	0.19
	D08 [verb+er-irreg]	0.55	0.48	0.41	0.91	0.86	0.79		0.30	0.24	0.17	0.12	0.10	0.07
	D09 [verb+tion-irreg]	0.63	0.46	0.41	0.91	0.93	0.86		0.51	0.22	0.15	0.61	0.61	0.51
	D10 [verb+ment-irreg]	0.64	0.46	0.42	0.92	0.90	0.90		0.47	0.24	0.15	0.54	0.50	0.44
	Encyclopedia	0.48	0.43	0.55	0.41	0.43	0.45		0.35	0.29	0.42	0.35	0.32	0.34
	E01 [country - capital]	0.63	0.47	0.41	0.96	0.96	0.98		0.61	0.35	0.32	0.90	0.90	0.90
	E02 [country - language]	0.40	0.43	0.59	0.33	0.24	0.33		0.36	0.31	0.45	0.43	0.21	0.36
	E03 [UK-city - county]	0.59	0.49	0.57	0.30	0.42	0.36		0.41	0.36	0.52	0.14	0.12	0.14
	E04 [name - nationality]	0.28	0.39	0.64	0.49	0.51	0.60		0.20	0.20	0.39	0.26	0.15	0.26
	E05 [name - occupation]	0.44	0.41	0.57	0.75	0.73	0.73		0.33	0.21	0.40	0.60	0.57	0.52
	E06 [animal - young]	0.47	0.43	0.44	0.10	0.15	0.15		0.34	0.36	0.38	0.09	0.12	0.12
	E07 [animal - sound]	0.37	0.43	0.40	0.22	0.17	0.22		0.15	0.31	0.25	0.08	0.00	0.00
	E08 [animal - shelter]	0.44	0.42	0.51	0.02	0.13	0.13		0.25	0.29	0.39	0.07	0.14	0.09
	E09 [things - color]	0.44	0.38	0.81	0.12	0.22	0.16		0.20	0.23	0.63	0.19	0.19	0.21
	E10 [male - female]	0.73	0.43	0.43	0.76	0.71	0.78		0.62	0.28	0.33	0.68	0.68	0.68
	Inflectional	0.75	0.48	0.51	0.97	0.97	0.97		0.63	0.22	0.23	0.88	0.88	0.88
	I01 [noun - plural-reg]	0.79	0.39	0.41	1.00	1.00	1.00		0.69	0.13	0.16	0.86	0.84	0.88
	I02 [noun - plural-irreg]	0.77	0.40	0.42	0.84	0.82	0.84		0.62	0.12	0.16	0.69	0.71	0.75
	I03 [adj - comparative]	0.75	0.50	0.52	1.00	1.00	1.00		0.63	0.23	0.37	0.97	1.00	1.00
	I04 [adj - superlative]	0.71	0.51	0.58	0.96	0.96	1.00		0.59	0.26	0.39	0.97	0.97	0.97
	I05 [verb-inf - 3pSg]	0.77	0.52	0.53	1.00	1.00	1.00		0.65	0.26	0.33	1.00	1.00	1.00
	I06 [verb-inf - Ving]	0.77	0.51	0.51	0.96	0.96	0.96		0.67	0.26	0.19	0.86	0.82	0.82
	I07 [verb-inf - Ved]	0.75	0.51	0.55	1.00	1.00	1.00		0.66	0.25	0.20	0.92	0.90	0.88
	I08 [verb-Ving - 3pSg]	0.70	0.48	0.51	0.98	0.98	0.98		0.56	0.17	0.31	0.90	0.90	0.90
	I09 [verb-Ving - Ved]	0.73	0.50	0.54	0.98	0.96	0.98		0.64	0.18	0.19	0.84	0.82	0.82
	I10 [verb-3pSg - Ved]	0.72	0.53	0.55	1.00	1.00	1.00		0.62	0.33	0.20	0.88	0.88	0.88
	Lexicography	0.62	0.37	0.38	0.21	0.22	0.26		0.45	0.17	0.18	0.19	0.17	0.18
	L01 [hypernyms - animals]	0.58	0.43	0.53	0.20	0.37	0.24		0.47	0.32	0.47	0.08	0.08	0.05
	L02 [hypernyms - misc]	0.55	0.35	0.39	0.23	0.16	0.14		0.42	0.21	0.21	0.26	0.26	0.10
	L03 [hyponyms - misc]	0.63	0.35	0.31	0.33	0.28	0.28		0.52	0.15	0.15	0.19	0.14	0.14
	L04 [meronyms - substance]	0.53	0.36	0.44	0.10	0.15	0.17		0.35	0.16	0.24	0.09	0.09	0.11
	L05 [meronyms - member]	0.58	0.36	0.36	0.12	0.10	0.12		0.36	0.15	0.15	0.10	0.06	0.08
	L06 [meronyms - part]	0.53	0.31	0.30	0.04	0.15	0.09		0.34	0.14	0.12	0.09	0.09	0.02
	L07 [synonyms - intensity]	0.67	0.37	0.37	0.27	0.32	0.36		0.51	0.17	0.16	0.26	0.26	0.30
	L08 [synonyms - exact]	0.71	0.33	0.31	0.18	0.16	0.22		0.55	0.11	0.11	0.15	0.15	0.22
	L09 [antonyms - gradable]	0.68	0.45	0.43	0.43	0.37	0.55		0.45	0.18	0.19	0.45	0.41	0.43
	L10 [antonyms - binary]	0.72	0.40	0.40	0.18	0.18	0.39		0.50	0.14	0.15	0.19	0.17	0.31
	Micro Mean	0.62	0.44	0.47	0.59	0.60	0.61		0.47	0.22	0.26	0.49	0.47	0.47
	Macro Mean	0.62	0.44	0.47	0.60	0.60	0.61		0.46	0.22	0.26	0.48	0.46	0.45
	Macro Std	0.12	0.06	0.09	0.34	0.33	0.32		0.14	0.07	0.12	0.32	0.32	0.32

Table 2: Detailed results for all combinations of FastText/Google News embeddings and Google Analogy and BATS analogies. In this table the cosine similarity is computed in the original space. See the main paper for more details.