

Supplemental Material

Global Constraints:

- **Single languages:**

af	am	ar	az	be	bg
bn	bs	ca	ceb	co	cs
cy	da	de	el	en	eo
es	et	eu	fa	fi	fil
fr	fy	ga	gd	gl	gu
ha	haw	he	hi	hi-Latn	hr
ht	hu	hy	id	ig	is
it	ja	jv	ka	kk	km
kn	ko	ku	ky	la	lb
lo	lt	lv	mg	mi	mk
ml	mn	mr	ms	mt	my
ne	nl	no	ny	or	pa
pl	ps	pt	ro	ru	si
sk	sl	sm	sn	so	sq
sr	su	sv	sw	ta	te
tg	th	tr	uk	ur	uz
vi	yi	yo	zh		

- **Language pairs:**

en/af	en/am	en/ar	en/az	en/be	en/bg
en/bn	en/bs	en/ca	en/ceb	en/co	en/cs
en/cy	en/da	en/de	en/el	en/eo	en/es
en/et	en/eu	en/fa	en/fi	en/fil	en/fr
en/fy	en/ga	en/gd	en/gl	en/gu	en/ha
en/haw	en/he	en/hi	en/hi-Latn	en/hr	en/ht
en/hu	en/hy	en/id	en/ig	en/is	en/it
en/ja	en/jv	en/ka	en/kk	en/km	en/kn
en/ko	en/ku	en/ky	en/la	en/lb	en/lo
en/lt	en/lv	en/mg	en/mi	en/mk	en/ml
en/mn	en/mr	en/ms	en/mt	en/my	en/ne
en/nl	en/no	en/ny	en/or	en/pa	en/pl
en/ps	en/pt	en/ro	en/ru	en/si	en/sk
en/sl	en/sm	en/sn	en/so	en/sq	en/sr
en/su	en/sv	en/sw	en/ta	en/te	en/tg
en/th	en/tr	en/uk	en/ur	en/uz	en/vi
en/yi	en/yo	en/zh	hi/hi-Latn		

Number of Parameters:

- Full model:

Variables	Size
Char unigram embedding	1000*16
Char bigram embedding	1000*16
Char trigram embedding	3000*16
Char quadgram embedding	5000*16
Script embedding	27*8
Lexicon language distribution embedding	16*100
Lexicon active languages embedding	16*100
Lexicon singletons embedding	16*100
Hidden layer W	$256*(16*3*7+8)$
Hidden layer bias	256
Softmax layer W	256*100
Softmax layer bias	100
<hr/>	
Total	279036

- Small model (without lexicon features):

Variables	Size
Char unigram embedding	1000*16
Char bigram embedding	1000*16
Char trigram embedding	3000*16
Char quadgram embedding	5000*16
Script embedding	27*8
Hidden layer W	$256*(16*3*4+8)$
Hidden layer bias	256

Softmax layer W	256*100
Softmax layer bias	100
<hr/>	
Total	237372

Per-language Results on KB-Mono56:

Test Count: Number of labels in prediction

Gold Count: Number of labels in gold

Label	F1-Score	Precision	Recall	Test Count	Gold Count
af	0.00	0.00	0.00	2	0
ar	99.84	99.79	99.89	1899	1897
be	0.00	0.00	0.00	3	0
bg	99.73	99.67	99.78	1842	1840
bn	100.00	100.00	100.00	2677	2677
bs	0.00	0.00	0.00	108	0
ca	95.44	99.96	91.31	2419	2648
ceb	0.00	0.00	0.00	3	0
co	0.00	0.00	0.00	12	0
cs	96.41	98.17	94.71	1588	1646
cy	97.25	100.00	94.64	2825	2985
da	94.51	100.00	89.58	1548	1728
de	97.49	100.00	95.10	2560	2692
el	100.00	100.00	100.00	2072	2072
en	57.34	40.30	99.34	6365	2582
eo	0.00	0.00	0.00	5	0
es	97.36	99.17	95.63	2515	2608
et	96.15	100.00	92.58	1235	1334

eu	96.74	99.78	93.87	1381	1468
fa	99.80	99.91	99.69	2235	2240
fi	96.46	100.00	93.15	1551	1665
fil	96.35	99.93	93.01	2757	2962
fr	97.15	99.07	95.29	2473	2571
fy	0.00	0.00	0.00	1	0
gl	0.00	0.00	0.00	3	0
gu	100.00	100.00	100.00	3647	3647
ha	0.00	0.00	0.00	4	0
hi	98.90	100.00	97.82	2916	2981
hi-Latn	0.00	0.00	0.00	1	0
hr	94.73	99.36	90.51	2045	2245
ht	95.59	100.00	91.55	1637	1788
hu	95.64	100.00	91.65	1909	2083
hy	100.00	100.00	100.00	1883	1883
id	58.35	42.99	90.79	4196	1987
is	97.93	99.90	96.04	2086	2170
it	95.67	99.67	91.97	1839	1993
iw	100.00	100.00	100.00	1829	1829
ja	99.77	100.00	99.55	3509	3525
ka	100.00	100.00	100.00	1649	1649
kn	100.00	100.00	100.00	2170	2170
ko	100.00	100.00	100.00	3362	3362
ku	0.00	0.00	0.00	9	0
lt	98.94	99.85	98.05	1360	1385

mi	0.00	0.00	0.00	1	0
mk	99.53	100.00	99.06	2098	2118
ml	100.00	100.00	100.00	1838	1838
mn	0.00	0.00	0.00	2	0
mr	99.88	99.86	99.90	2086	2085
ms	23.23	88.21	13.37	458	3021
ne	98.72	97.46	100.00	2445	2383
nl	97.05	99.96	94.30	2350	2491
no	95.91	97.04	94.80	2198	2250
or	100.00	100.00	100.00	5534	5534
pa	100.00	100.00	100.00	3930	3930
pl	97.47	99.95	95.10	1905	2002
ps	0.00	0.00	0.00	2	0
pt	97.63	100.00	95.37	3048	3196
ro	96.29	99.96	92.88	2455	2642
ru	99.75	100.00	99.51	1423	1430
si	100.00	100.00	100.00	2499	2499
sk	93.07	99.89	87.11	1807	2072
sl	95.32	100.00	91.06	1629	1789
sr	0.00	0.00	0.00	20	0
sv	97.38	99.84	95.03	1935	2033
sw	0.00	0.00	0.00	13	0
ta	100.00	100.00	100.00	1855	1855
te	100.00	100.00	100.00	3470	3470
th	100.00	100.00	100.00	3253	3253

tr	98.55	99.55	97.57	1573	1605
uk	100.00	100.00	100.00	1595	1595
vi	97.48	100.00	95.09	3274	3443
yo	0.00	0.00	0.00	8	0
zh	99.89	99.78	100.00	5530	5518

Language Pairs for Synthetic Data:

en/af	en/am	en/ar	en/az	en/be	en/bg
en/bn	en/bs	en/ca	en/ceb	en/co	en/cs
en/cy	en/da	en/de	en/el	en/eo	en/es
en/et	en/eu	en/fa	en/fi	en/fil	en/fr
en/fy	en/ga	en/gd	en/gl	en/gu	en/ha
en/haw	en/he	en/hi	en/hi-Latn	en/hr	en/ht
en/hu	en/hy	en/id	en/ig	en/is	en/it
en/ja	en/jv	en/ka	en/kk	en/km	en/kn
en/ko	en/ku	en/ky	en/la	en/lb	en/lo
en/lt	en/lv	en/mg	en/mi	en/mk	en/ml
en/mn	en/mr	en/ms	en/mt	en/my	en/ne
en/nl	en/no	en/ny	en/or	en/pa	en/pl
en/ps	en/pt	en/ro	en/ru	en/si	en/sk
en/sl	en/sm	en/sn	en/so	en/sq	en/sr
en/su	en/sv	en/sw	en/ta	en/te	en/tg
en/th	en/tr	en/uk	en/ur	en/uz	en/vi
en/yi	en/yo	en/zh	hi/hi-Latn		