

EWT (basic)

trained on: en_ewt-ud-train.iob2 + en_ewt-ud-dev.iob2

tested on: lyrics_test.iob2

recall: 0.7777076097252921

precision: 0.8295722465476592

slot-f1: 0.8028031290743154

unlabeled

ul_recall: 0.8190716766656142

ul_precision: 0.8736948467497474

ul_slot-f1: 0.8455019556714474

loose (partial overlap with same label)

l_recall: 0.8288601199873697

l_precision: 0.8858201414617717

l_slot-f1: 0.8563940522867102

Mistake breakdown:

Correct 29740

Wrong_label 307

Missed 791

Spurious 256

I-PER->B-PER 56

B-ORG->B-PER 41

B-PER->B-LOC 40

B-ORG->B-LOC 35

I-ORG->I-PER 22

I-ORG->I-LOC 20

B-LOC->B-PER 16

I-PER->I-LOC 15

B-PER->B-ORG 13

I-LOC->B-LOC 9

I-PER->I-ORG 8

B-PER->I-PER 8

B-LOC->B-ORG 6

I-LOC->I-PER 6

I-ORG->B-PER 3

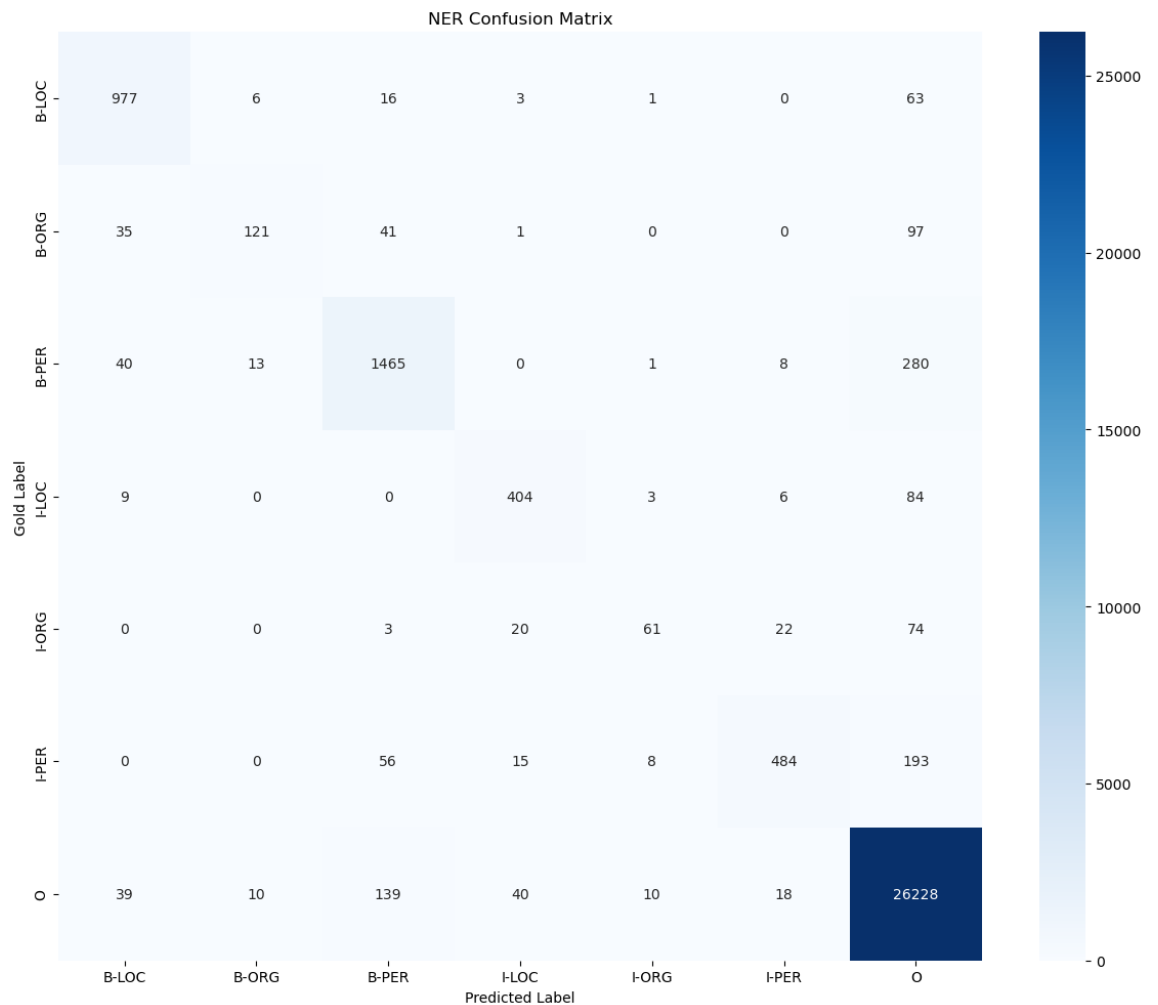
B-LOC->I-LOC 3

I-LOC->I-ORG 3

B-LOC->I-ORG 1

B-ORG->I-LOC 1

B-PER->I-ORG 1



EWT (DAPT)

trained on: en_ewt-ud-train.iob2 + en_ewt-ud-dev.iob2

tested on: lyrics_test.iob2

recall: 0.7859172718661194

precision: 0.8434429007116232

slot-f1: 0.813664596273292

unlabeled

ul_recall: 0.8228607514998422

ul_precision: 0.8830904778041342

ul_slot-f1: 0.8519123896698266

loose (partial overlap with same label)

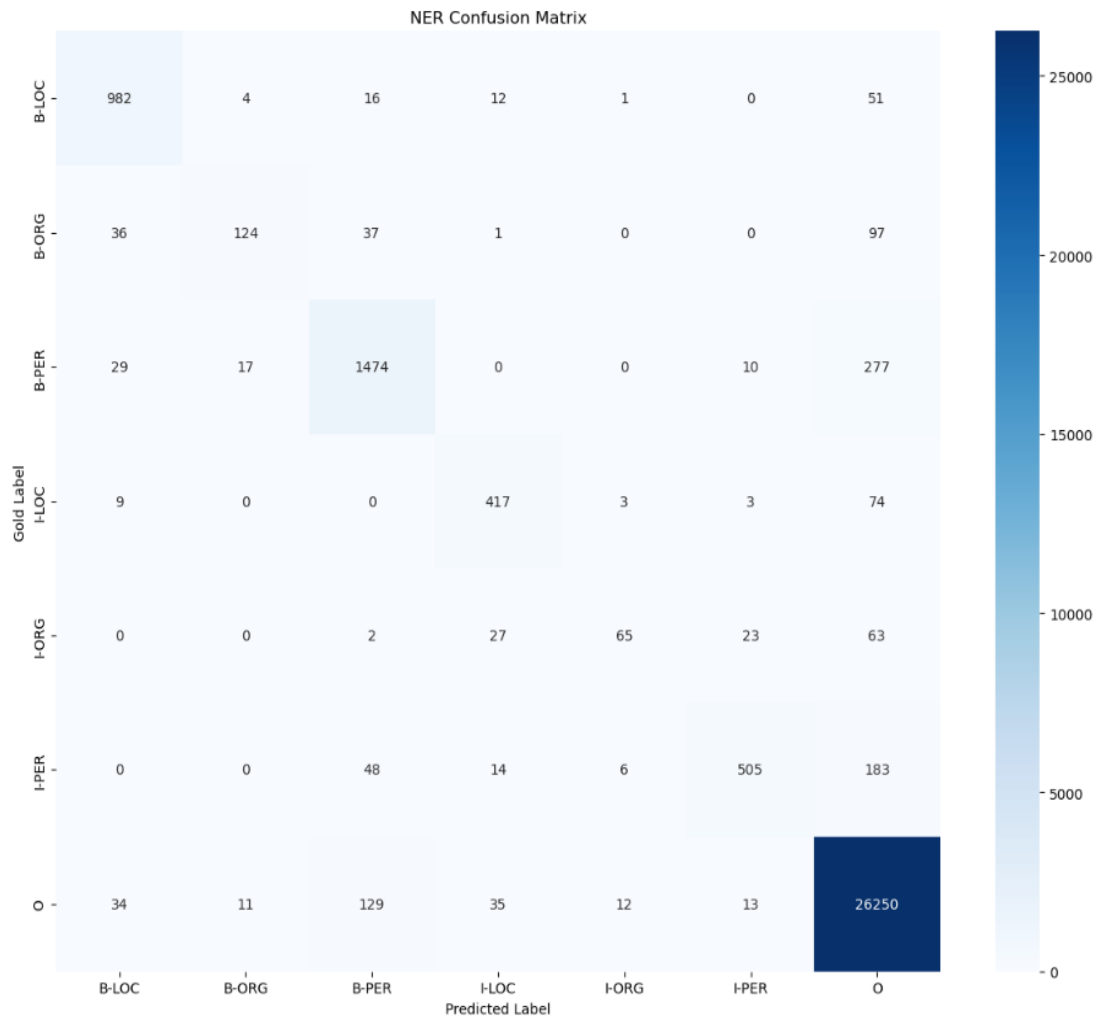
l_recall: 0.8320176823492264

l_precision: 0.8952897322941376

l_slot-f1: 0.8624948654530336

Mistake breakdown:

Correct 29817
Wrong_label 298
Missed 745
Spurious 234
I-PER->B-PER 48
B-ORG->B-PER 37
B-ORG->B-LOC 36
B-PER->B-LOC 29
I-ORG->I-LOC 27
I-ORG->I-PER 23
B-PER->B-ORG 17
B-LOC->B-PER 16
I-PER->I-LOC 14
B-LOC->I-LOC 12
B-PER->I-PER 10
I-LOC->B-LOC 9
I-PER->I-ORG 6
B-LOC->B-ORG 4
I-LOC->I-ORG 3
I-LOC->I-PER 3
I-ORG->B-PER 2
B-LOC->I-ORG 1
B-ORG->I-LOC 1



EWT (DAPT) + pop

trained on: en_ewt-ud-train.iob2 + en_ewt-ud-dev.iob2 + pop-manual-1000-train.iob2

tested on: lyrics_test.iob2

recall: 0.8575939374802652

precision: 0.8442648430214486

slot-f1: 0.8508771929824562

unlabeled

ul_recall: 0.9081149352699716

ul_precision: 0.8940006216972335

ul_slot-f1: 0.9010025062656642

loose (partial overlap with same label)

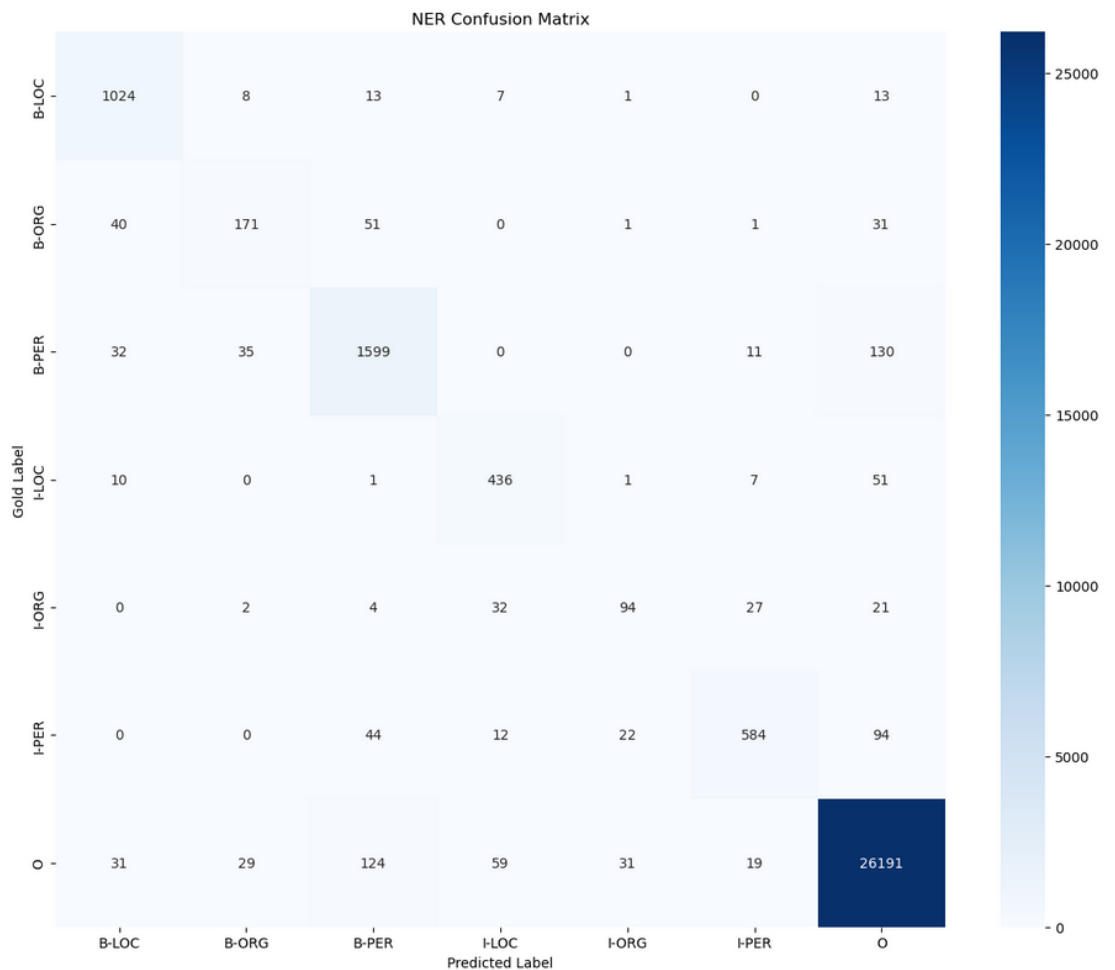
l_recall: 0.897379223239659

l_precision: 0.887161952129313

l_slot-f1: 0.8922413385333924

Mistake breakdown:

Correct 30099
Wrong_label 362
Missed 340
Spurious 293
B-ORG->B-PER 51
I-PER->B-PER 44
B-ORG->B-LOC 40
B-PER->B-ORG 35
B-PER->B-LOC 32
I-ORG->I-LOC 32
I-ORG->I-PER 27
I-PER->I-ORG 22
B-LOC->B-PER 13
I-PER->I-LOC 12
B-PER->I-PER 11
I-LOC->B-LOC 10
B-LOC->B-ORG 8
B-LOC->I-LOC 7
I-LOC->I-PER 7
I-ORG->B-PER 4
I-ORG->B-ORG 2
I-LOC->I-ORG 1
B-LOC->I-ORG 1
B-ORG->I-ORG 1
I-LOC->B-PER 1
B-ORG->I-PER 1



EWT (DAPT) + pop (cont.)

trained on: en_ewt-ud-train.iob2 + en_ewt-ud-dev.iob2 + pop-manual-1000-train.iob2 + pop_labeled_no_2000.iob2
tested on: lyrics_test.iob2

recall: 0.8497000315756236

precision: 0.8605692356891589

slot-f1: 0.8551000953288845

unlabeled

ul_recall: 0.8958004420587307

ul_precision: 0.9072593540134314

ul_slot-f1: 0.9014934858595487

loose (partial overlap with same label)

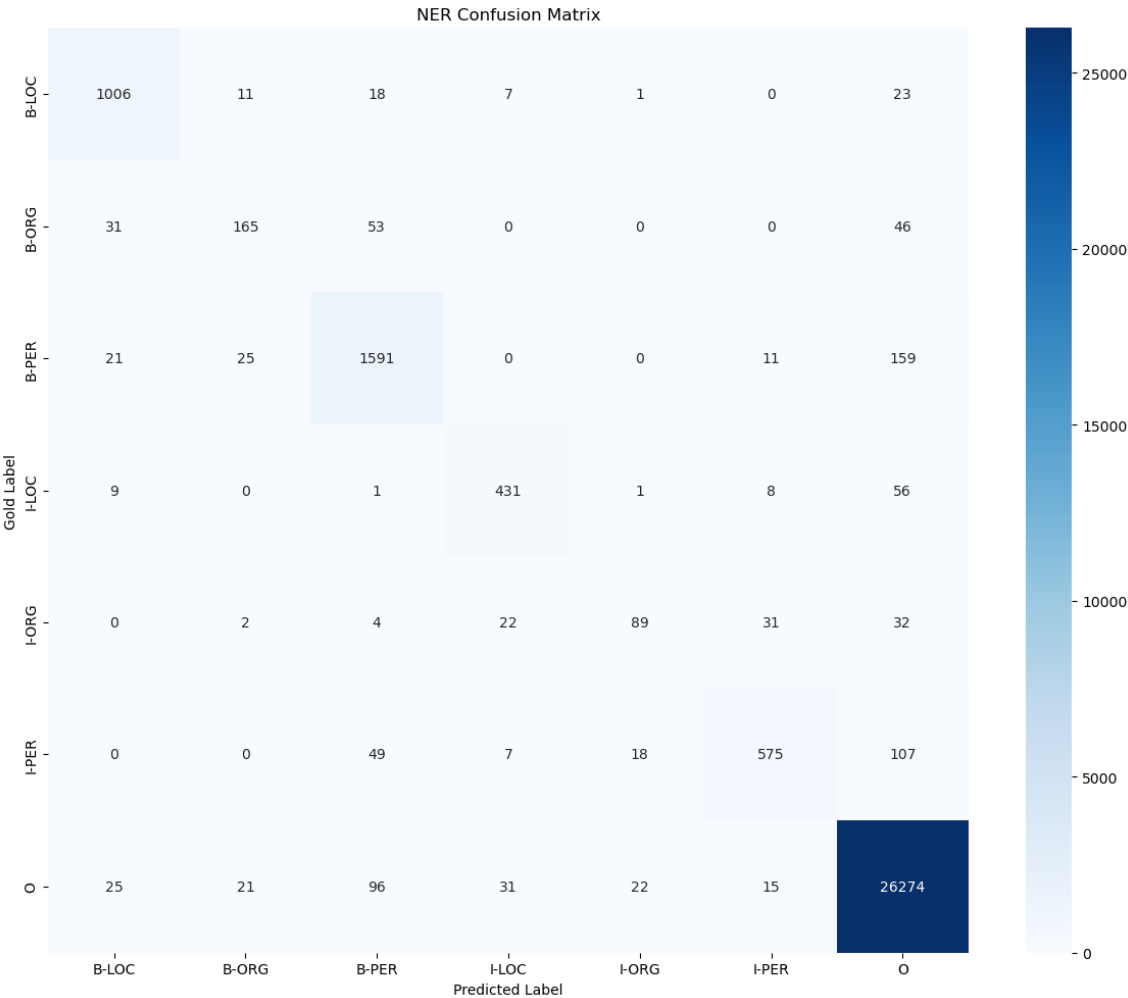
l_recall: 0.8866435112093464

l_precision: 0.9021426287176207

l_slot-f1: 0.8943259231318492

Mistake breakdown:

- Correct 30131
- Wrong_label 330
- Missed 423
- Spurious 210
- B-ORG->B-PER 53
- I-PER->B-PER 49
- I-ORG->I-PER 31
- B-ORG->B-LOC 31
- B-PER->B-ORG 25
- I-ORG->I-LOC 22
- B-PER->B-LOC 21
- B-LOC->B-PER 18
- I-PER->I-ORG 18
- B-LOC->B-ORG 11
- B-PER->I-PER 11
- I-LOC->B-LOC 9
- I-LOC->I-PER 8
- I-PER->I-LOC 7
- B-LOC->I-LOC 7
- I-ORG->B-PER 4
- I-ORG->B-ORG 2
- I-LOC->I-ORG 1
- B-LOC->I-ORG 1
- I-LOC->B-PER 1



EWT (DAPT) + country

trained on: en_ewt-ud-train.iob2 + en_ewt-ud-dev.iob2 + country-manual-1000-train.iob2

tested on: lyrics_test.iob2

recall: 0.8680138932743922

precision: 0.8260216346153846

slot-f1: 0.8464973056197074

unlabeled

ul_recall: 0.9163245974107989

ul_precision: 0.8719951923076923

ul_slot-f1: 0.8936104695919939

loose (partial overlap with same label)

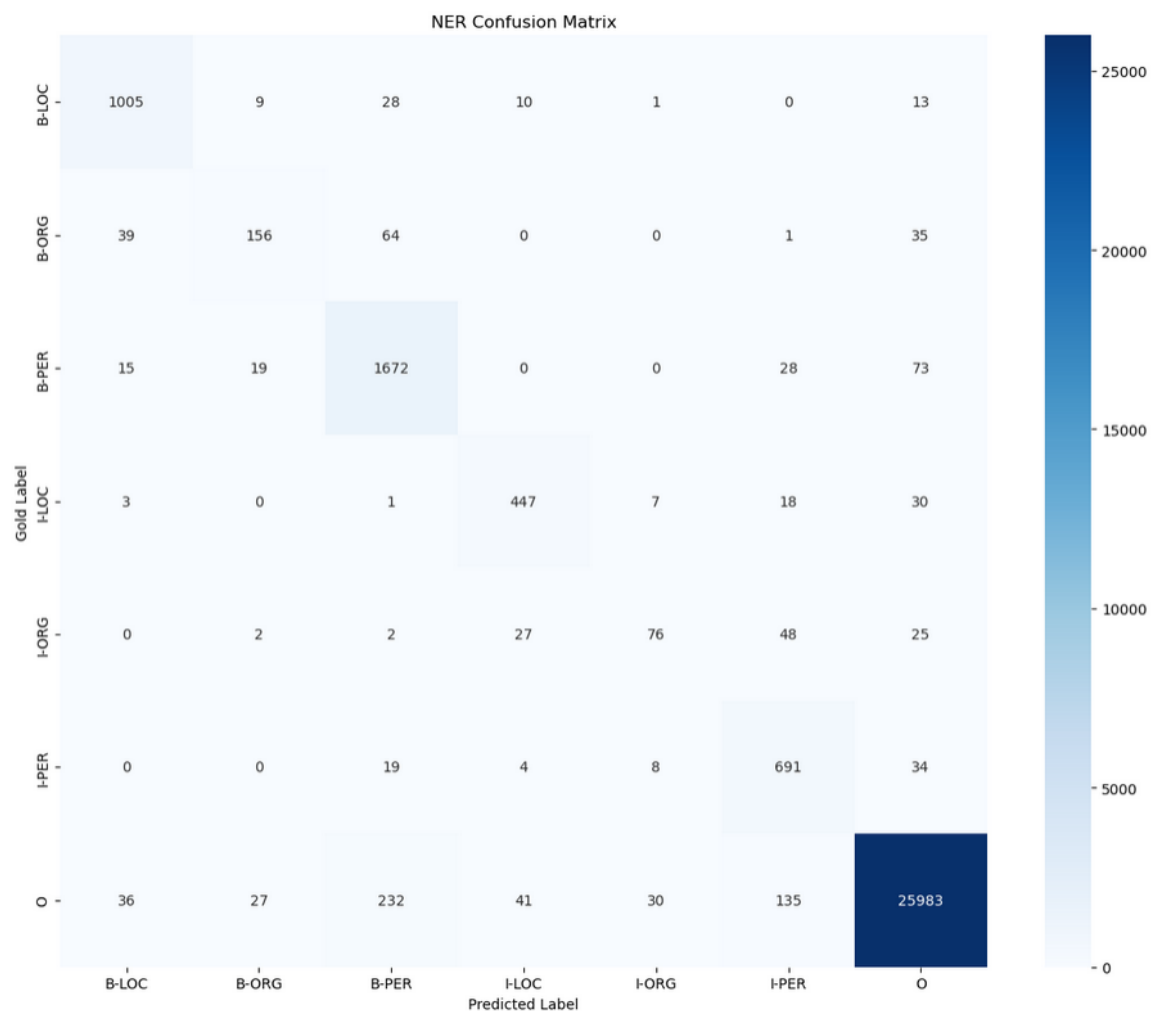
l_recall: 0.905904641616672

l_precision: 0.8620793269230769

l_slot-f1: 0.8834488067744418

Mistake breakdown:

Correct 30030
Wrong_label 353
Missed 210
Spurious 501
B-ORG->B-PER 64
I-ORG->I-PER 48
B-ORG->B-LOC 39
B-LOC->B-PER 28
B-PER->I-PER 28
I-ORG->I-LOC 27
I-PER->B-PER 19
B-PER->B-ORG 19
I-LOC->I-PER 18
B-PER->B-LOC 15
B-LOC->I-LOC 10
B-LOC->B-ORG 9
I-PER->I-ORG 8
I-LOC->I-ORG 7
I-PER->I-LOC 4
I-LOC->B-LOC 3
I-ORG->B-PER 2
I-ORG->B-ORG 2
B-LOC->I-ORG 1
B-ORG->I-PER 1
I-LOC->B-PER 1



EWT (DAPT) + country (cont.)

trained on: en_ewt-ud-train.iob2 + en_ewt-ud-dev.iob2 + country-manual-1000-train.iob2 + country_labeled_no_2000.iob2
tested on: lyrics_test.iob2

recall: 0.8572781812440796

precision: 0.8447417548226509

slot-f1: 0.8509637987776212

unlabeled

ul_recall: 0.904957372908115

ul_precision: 0.8917237087741132

ul_slot-f1: 0.8982918037925091

loose (partial overlap with same label)

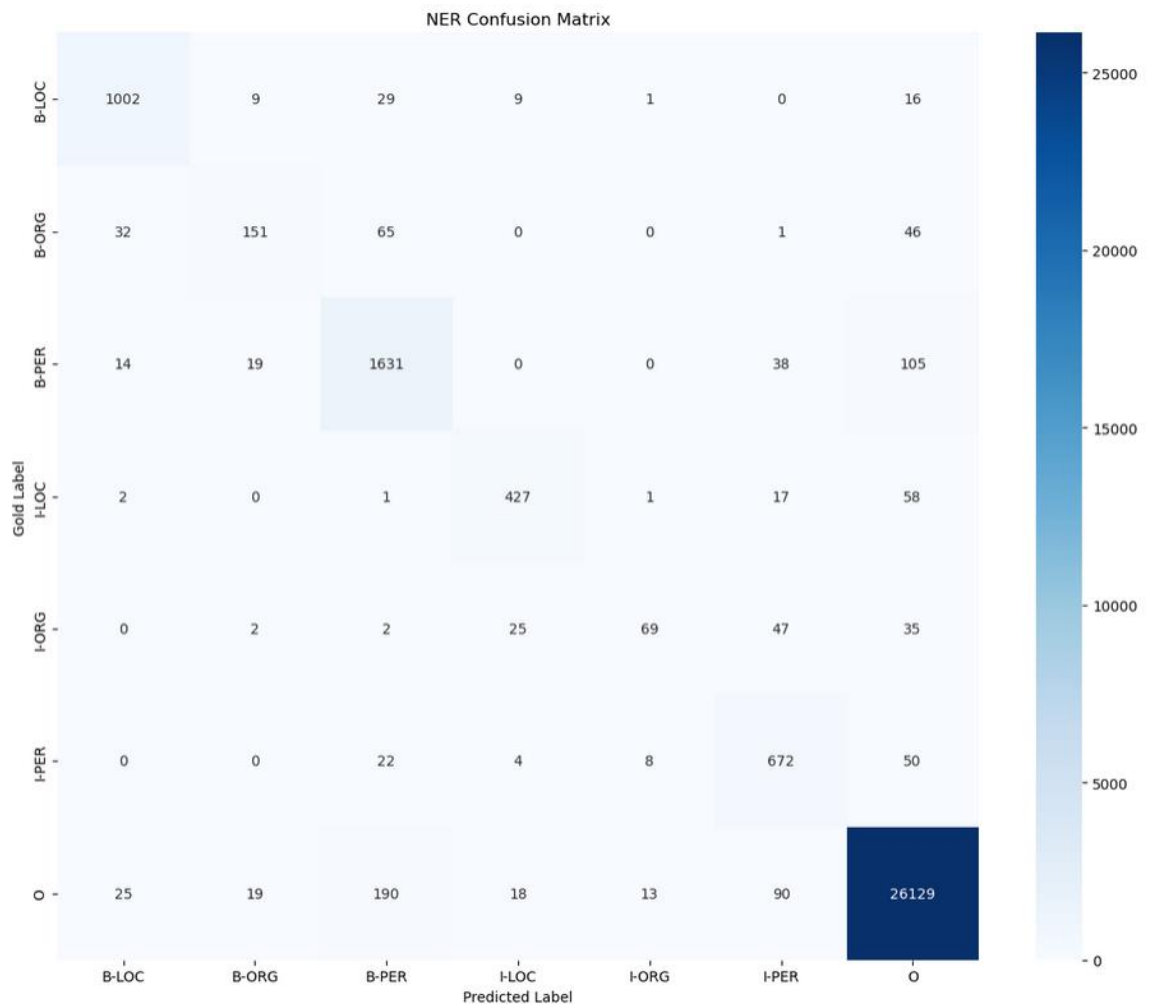
l_recall: 0.8901168298073887

l_precision: 0.8755444928438083

l_slot-f1: 0.8827705271985561

Mistake breakdown:

Correct 30081
Wrong_label 348
Missed 310
Spurious 355
B-ORG->B-PER 65
I-ORG->I-PER 47
B-PER->I-PER 38
B-ORG->B-LOC 32
B-LOC->B-PER 29
I-ORG->I-LOC 25
I-PER->B-PER 22
B-PER->B-ORG 19
I-LOC->I-PER 17
B-PER->B-LOC 14
B-LOC->I-LOC 9
B-LOC->B-ORG 9
I-PER->I-ORG 8
I-PER->I-LOC 4
I-LOC->B-LOC 2
I-ORG->B-PER 2
I-ORG->B-ORG 2
B-LOC->I-ORG 1
I-LOC->I-ORG 1
B-ORG->I-PER 1
I-LOC->B-PER 1



EWT (DAPT) + rap/hip-hop

trained on: en_ewt-ud-train.iob2 + en_ewt-ud-dev.iob2 + rap-hip-hop-manual-1000-train.iob2
tested on: lyrics_test.iob2

recall: 0.8711714556362489
precision: 0.8335347432024169
slot-f1: 0.8519376254438784

unlabeled

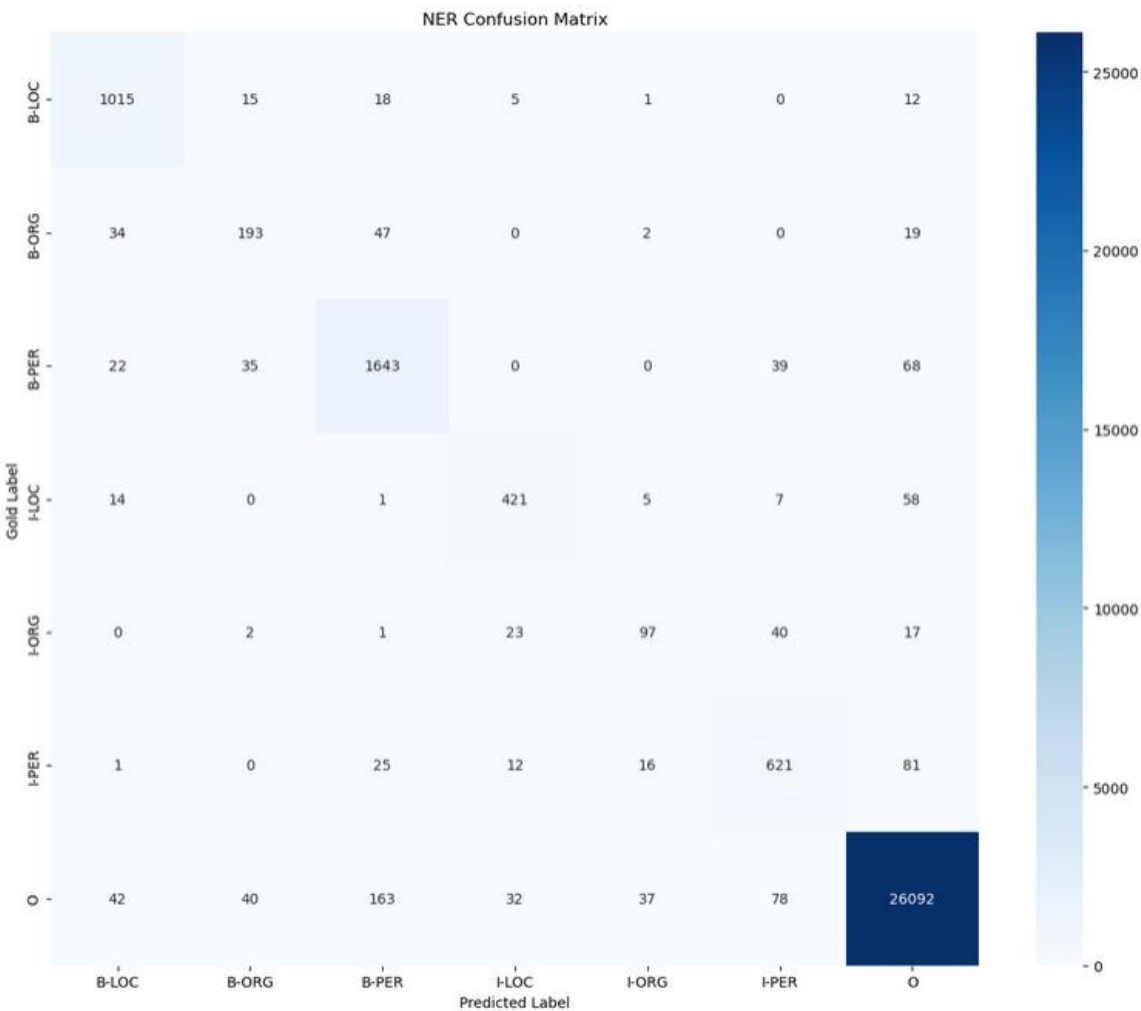
ul_recall: 0.9216924534259552
ul_precision: 0.8818731117824773
ul_slot-f1: 0.9013432144511347

loose (partial overlap with same label)

l_recall: 0.9172718661193558
l_precision: 0.8824773413897281
l_slot-f1: 0.8995382627393301

Mistake breakdown:

Correct 30082
Wrong_label 365
Missed 255
Spurious 392
B-ORG->B-PER 47
I-ORG->I-PER 40
B-PER->I-PER 39
B-PER->B-ORG 35
B-ORG->B-LOC 34
I-PER->B-PER 25
I-ORG->I-LOC 23
B-PER->B-LOC 22
B-LOC->B-PER 18
I-PER->I-ORG 16
B-LOC->B-ORG 15
I-LOC->B-LOC 14
I-PER->I-LOC 12
I-LOC->I-PER 7
B-LOC->I-LOC 5
I-LOC->I-ORG 5
B-ORG->I-ORG 2
I-ORG->B-ORG 2
B-LOC->I-ORG 1
I-PER->B-LOC 1
I-LOC->B-PER 1
I-ORG->B-PER 1



EWT (DAPT) + rap/hip-hop (cont.)

trained on: en_ewt-ud-train.iob2 + en_ewt-ud-dev.iob2 + rap-hip-hop-manual-1000-train.iob2 + rap-hip-hop_labeled_no_2000.iob2

tested on: lyrics_test.iob2

recall: 0.8746447742342911

precision: 0.8429701765063907

slot-f1: 0.8585154191848752

unlabeled

ul_recall: 0.9210609409535838

ul_precision: 0.8877054169202678

ul_slot-f1: 0.9040756237408957

loose (partial overlap with same label)

l_recall: 0.9204294284812125

l_precision: 0.8919659160073037

l_slot-f1: 0.9059741637407696

Mistake breakdown:

Correct 30136

Wrong_label :344

Missed 254

Spurious 360

B-ORG->B-PER 44

B-PER->I-PER 41

I-ORG->I-PER 36

B-PER->B-ORG 31

B-ORG->B-LOC 29

I-ORG->I-LOC 23

I-PER->B-PER 23

B-PER->B-LOC 20

B-LOC->B-PER 18

I-LOC->B-LOC 16

I-PER->I-LOC 15

B-LOC->B-ORG 13

I-PER->I-ORG 12

I-LOC->I-PER 9

B-LOC->I-LOC 4

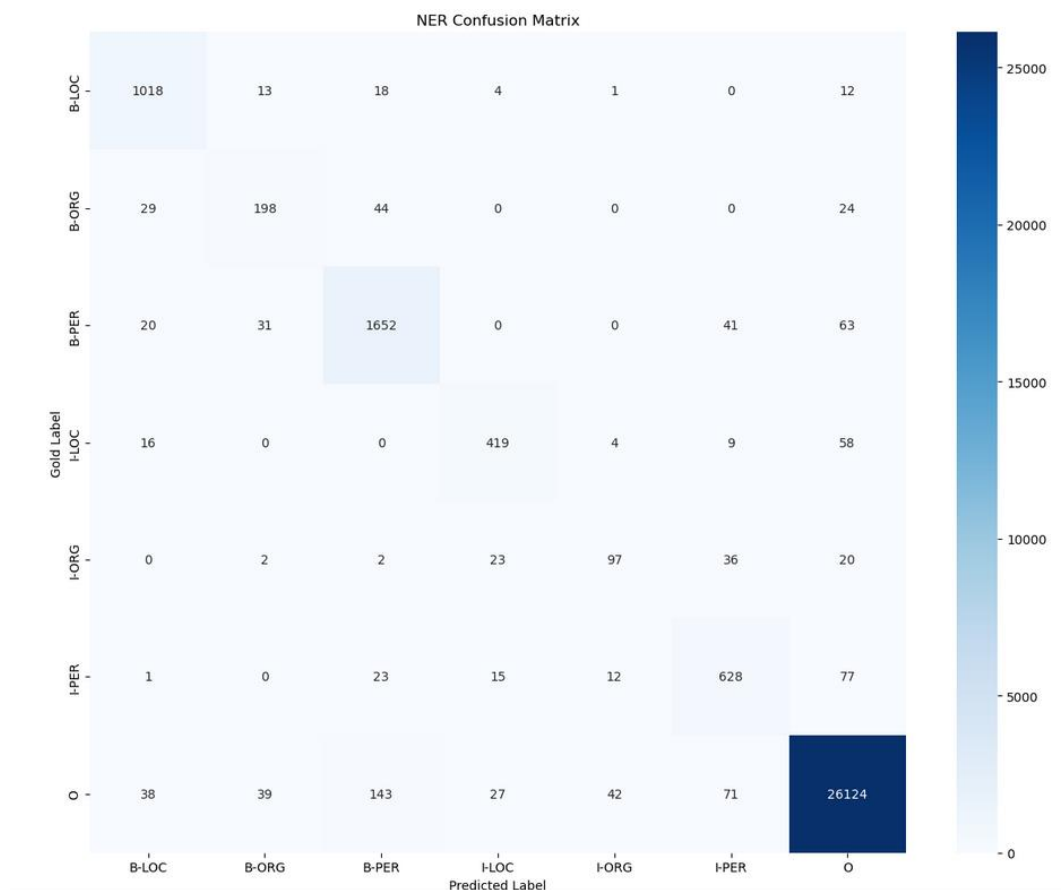
I-LOC->I-ORG 4

I-ORG->B-PER 2

I-ORG->B-ORG 2

B-LOC->I-ORG 1

I-PER->B-LOC 1



EWT (DAPT) + all genres

trained on: en_ewt-ud-train.iob2 + en_ewt-ud-dev.iob2 + pop-manual-1000-train.iob2 + country-manual-1000-train.iob2 + rap-hip-hop-manual-1000-train.iob2

tested on: lyrics_test.iob2

recall: 0.8853804862646037

precision: 0.8640986132511557

slot-f1: 0.8746101060511541

unlabeled

ul_recall: 0.9305336280391537

ul_precision: 0.9081664098613251

ul_slot-f1: 0.9192139737991265

loose (partial overlap with same label)

l_recall: 0.9210609409535838

l_precision: 0.8983050847457628

l_slot-f1: 0.9095407025656412

Mistake breakdown:

Correct 30270

Wrong_label 308

Missed 220

Spurious 296

B-ORG->B-PER 49

I-ORG->I-PER 34

B-ORG->B-LOC 33

B-PER->B-ORG 27

B-PER->I-PER 26

I-PER->B-PER 24

I-ORG->I-LOC 24

B-LOC->B-PER 18

B-PER->B-LOC 16

I-PER->I-ORG 12

I-LOC->B-LOC 9

B-LOC->B-ORG 9

I-LOC->I-PER 9

I-PER->I-LOC 6

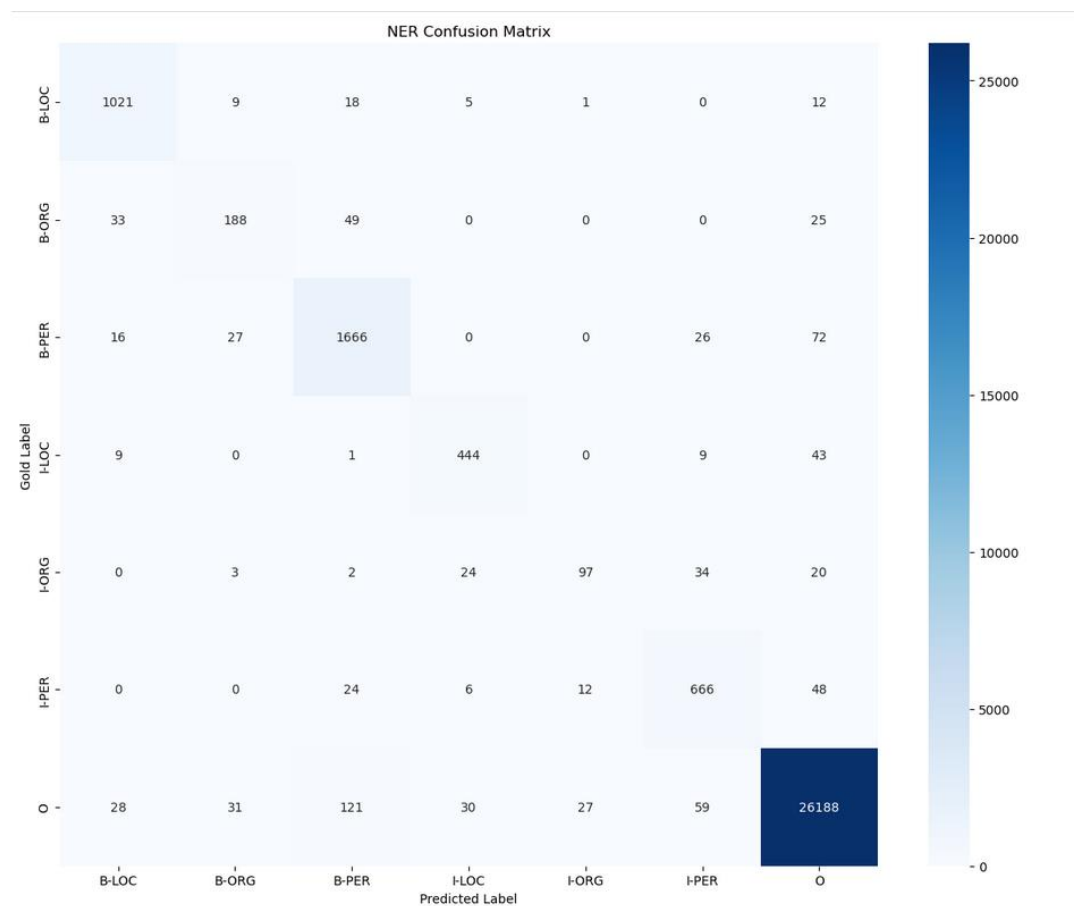
B-LOC->I-LOC 5

I-ORG->B-ORG 3

I-ORG->B-PER 2

B-LOC->I-ORG 1

I-LOC->B-PER 1



EWT (DAPT) + all genres (cont.)

trained on: en_ewt-ud-train.iob2 + en_ewt-ud-dev.iob2 + pop-manual-1000-train.iob2 + pop_labeled_no_2000.iob2 + country-manual-1000-train.iob2 + country_labeled_no_2000.iob2 + rap-hip-hop-manual-1000-train.iob2 + rap-hip-hop_labeled_no_2000.iob2

tested on: lyrics_test.iob2

recall: 0.8828544363751184

precision: 0.8667079975201488

slot-f1: 0.8747067104645706

unlabeled

ul_recall: 0.9251657720239975

ul_precision: 0.9082455052696838

ul_slot-f1: 0.9166275613952761

loose (partial overlap with same label)

l_recall: 0.9216924534259552

l_precision: 0.9070055796652201

l_slot-f1: 0.9142900389951772

Mistake breakdown:

Correct 30260
Wrong_label 309
Missed 221
Spurious 304
B-ORG->B-PER 44
I-ORG->I-PER 35
B-ORG->B-LOC 30
B-PER->I-PER 29
I-ORG->I-LOC 28
B-PER->B-ORG 24
I-PER->B-PER 21
B-LOC->B-PER 18
B-PER->B-LOC 15
I-LOC->B-LOC 13
B-LOC->B-ORG 13
I-LOC->I-PER 10
I-PER->I-ORG 10
I-PER->I-LOC 7
I-ORG->B-ORG 4
B-LOC->I-LOC 3
I-ORG->B-PER 2
B-PER->I-ORG 2
I-LOC->B-PER 1

