

Experimento: EX3 - Etapa 1 (SciERC e Material Science)

Diário de bordo com os ajustes finos e valores dos hiperparâmetros utilizados

Resultados do treinamento (ajuste fino dos modelos pré-treinados)

Corpus: -Corpus-1-SCIERC-Com uma só categoria de entidade-NER

Data: 16/11/2023

Parâmetros diferentes:

i) Notebook: IDEA-ETAPA 3-FineTuneBERT_Spacy_NER.ipynb

Vocabulário: en_core_web_sm

Corpus: SCIERC

ii) Arquivo de configuração: config_spacy.cfg

[nlp]

Linha 12: lang = "en"

[components.transformer.model]

Linha 52: name = "allenai/scibert_scivocab_uncased"

[training]

dropout = 0.2

iii) Resultados

E	#	LOSS TRANS...	LOSS NER	ENTS_F	ENTS_P	ENTS_R	SCORE
0	0	1201.91	209.26	3.79	2.71	6.29	0.04
28	200	56398.77	69224.21	76.68	74.75	78.72	0.77
57	400	335.41	31870.74	76.22	73.74	78.87	0.76
85	600	155.28	30465.57	76.44	72.72	80.56	0.76
114	800	59.64	29153.56	76.40	74.35	78.56	0.76
142	1000	108.42	28015.72	76.01	75.65	76.38	0.76
171	1200	114.84	26486.11	75.89	74.87	76.94	0.76
200	1400	49.35	24893.30	76.24	75.76	76.72	0.76
228	1600	85.12	23137.14	76.66	76.63	76.70	0.77
257	1800	40.50	21114.88	76.18	75.22	77.17	0.76

TOK 100.00

NER P 76.67
NER R 79.13
NER F 77.88
SPEED 3404

Corpus: -Corpus-1-SCIERC-Com uma só categoria de entidade-RE
Data: 16/11/2023

Parâmetros diferentes:

i) Notebook: IDEA-ETAPA 3-FineTuneBERT_Spacy_RE.ipynb

Vocabulário: en_core_web_sm
Corpus: Materiais químicos (RE)

ii) Arquivo de configuração: \configs\config_spacy.cfg

Linha 12: lang = "en"
Linha 19: batch_size = 500
Linha 30: name = "allenai/scibert_scivocab_uncased"
Linha 40: threshold = 0.5
Linha 60: max_length = 30
Linha 84: dropout = 0.2

Resultados do treinamento

E	#	LOSS TRANS...	LOSS RELAT...	REL_MICRO_P	REL_MICRO_R	REL_MICRO_F	SCORE
---	---	---------------	---------------	-------------	-------------	-------------	-------

0	0	0.82	4.47	0.82	46.50	1.61	0.02
14	100	111.66	88.96	52.99	2.06	3.96	0.04
28	200	1.06	9.05	49.72	14.68	22.67	0.23
42	300	0.60	5.89	42.14	24.14	30.70	0.31
57	400	0.25	4.15	48.67	19.50	27.84	0.28
71	500	0.06	3.36	48.46	20.39	28.71	0.29
85	600	0.05	3.23	52.19	18.60	27.42	0.27
100	700	0.04	3.19	48.46	20.86	29.16	0.29
114	800	0.01	2.98	49.39	21.62	30.08	0.30
128	900	0.01	2.97	50.65	20.56	29.25	0.29
142	1000	0.01	2.86	48.94	20.72	29.12	0.29

Random baseline:

threshold 0.00 {'rel_micro_p': '0.90', 'rel_micro_r': '100.00', 'rel_micro_f': '1.78'}
threshold 0.05 {'rel_micro_p': '0.90', 'rel_micro_r': '94.65', 'rel_micro_f': '1.77'}

threshold 0.10	{'rel_micro_p': '0.89', 'rel_micro_r': '89.19', 'rel_micro_f': '1.77'}
threshold 0.20	{'rel_micro_p': '0.90', 'rel_micro_r': '79.85', 'rel_micro_f': '1.78'}
threshold 0.30	{'rel_micro_p': '0.88', 'rel_micro_r': '68.84', 'rel_micro_f': '1.75'}
threshold 0.40	{'rel_micro_p': '0.87', 'rel_micro_r': '58.45', 'rel_micro_f': '1.72'}
threshold 0.50	{'rel_micro_p': '0.88', 'rel_micro_r': '49.21', 'rel_micro_f': '1.74'}
threshold 0.60	{'rel_micro_p': '0.86', 'rel_micro_r': '38.41', 'rel_micro_f': '1.68'}
threshold 0.70	{'rel_micro_p': '0.85', 'rel_micro_r': '28.44', 'rel_micro_f': '1.64'}
threshold 0.80	{'rel_micro_p': '0.80', 'rel_micro_r': '17.73', 'rel_micro_f': '1.52'}
threshold 0.90	{'rel_micro_p': '0.77', 'rel_micro_r': '8.60', 'rel_micro_f': '1.42'}
threshold 0.99	{'rel_micro_p': '0.58', 'rel_micro_r': '0.63', 'rel_micro_f': '0.61'}
threshold 1.00	{'rel_micro_p': '2.30', 'rel_micro_r': '0.21', 'rel_micro_f': '0.38'}

Results of the trained model:

threshold 0.00	{'rel_micro_p': '0.90', 'rel_micro_r': '100.00', 'rel_micro_f': '1.78'}
threshold 0.05	{'rel_micro_p': '20.92', 'rel_micro_r': '54.77', 'rel_micro_f': '30.28'}
threshold 0.10	{'rel_micro_p': '26.49', 'rel_micro_r': '47.53', 'rel_micro_f': '34.02'}
threshold 0.20	{'rel_micro_p': '32.98', 'rel_micro_r': '38.82', 'rel_micro_f': '35.66'}
threshold 0.30	{'rel_micro_p': '37.00', 'rel_micro_r': '33.89', 'rel_micro_f': '35.38'}
threshold 0.40	{'rel_micro_p': '40.52', 'rel_micro_r': '29.17', 'rel_micro_f': '33.92'}
threshold 0.50	{'rel_micro_p': '43.68', 'rel_micro_r': '26.13', 'rel_micro_f': '32.70'}
threshold 0.60	{'rel_micro_p': '47.20', 'rel_micro_r': '22.98', 'rel_micro_f': '30.91'}
threshold 0.70	{'rel_micro_p': '50.29', 'rel_micro_r': '18.36', 'rel_micro_f': '26.90'}
threshold 0.80	{'rel_micro_p': '52.78', 'rel_micro_r': '13.96', 'rel_micro_f': '22.07'}
threshold 0.90	{'rel_micro_p': '52.45', 'rel_micro_r': '7.87', 'rel_micro_f': '13.69'}
threshold 0.99	{'rel_micro_p': '37.50', 'rel_micro_r': '0.63', 'rel_micro_f': '1.24'}
threshold 1.00	{'rel_micro_p': '0.00', 'rel_micro_r': '0.00', 'rel_micro_f': '0.00'}

Corpus: -Corpus-2-Quimico-Com uma só categoria de entidade-NER

Data: 15/11/2023

Parâmetros diferentes:

i) Notebook: IDEA-ETAPA 3-FineTuneBERT_Spacy_NER.ipynb

Vocabulário: en_core_web_sm

ii) Arquivo de configuração: config_spacy.cfg

[nlp]

lang = "en"

[components.transformer.model]

name = "roberta-base"

[training]

dropout = 0.2

iii) Resultados

E	#	LOSS	TRANS...	LOSS NER	ENTS_F	ENTS_P	ENTS_R	SCORE
---	---	------	----------	----------	--------	--------	--------	-------

0	0	8908.69	533.21	0.75	1.61	0.49	0.01	
18	200	248616.74	74047.30	75.76	75.55	75.97	0.76	
36	400	3524.53	4508.36	78.06	78.31	77.81	0.78	
54	600	572.52	683.98	78.60	76.28	81.07	0.79	
72	800	364.21	453.76	78.71	76.74	80.78	0.79	
90	1000	336.83	360.10	78.59	78.17	79.01	0.79	
109	1200	243.87	251.69	79.56	75.72	83.79	0.80	
127	1400	122.40	130.01	79.70	78.99	80.42	0.80	
145	1600	81.82	85.97	79.45	76.97	82.09	0.79	
163	1800	93.08	88.65	79.74	78.08	81.46	0.80	
181	2000	59.50	63.58	79.72	77.03	82.60	0.80	
200	2200	66.49	63.73	79.55	76.22	83.18	0.80	
218	2400	46.63	43.53	79.29	76.96	81.77	0.79	
236	2600	42.58	49.92	80.15	80.91	79.41	0.80	
254	2800	153.48	144.06	79.62	77.80	81.53	0.80	
272	3000	80.79	66.17	79.66	77.67	81.76	0.80	
290	3200	52.27	46.44	79.51	77.48	81.64	0.80	
309	3400	78.58	66.07	79.68	77.89	81.55	0.80	
327	3600	43.39	31.21	79.82	77.96	81.77	0.80	
345	3800	74.52	58.21	79.74	78.72	80.79	0.80	
363	4000	29.70	14.96	79.10	75.81	82.69	0.79	
381	4200	67.01	53.91	79.98	77.48	82.64	0.80	

TOK 100.00
NER P 79.46
NER R 75.46
NER F 77.41
SPEED 6173

Corpus: -Corpus-2-Quimico-Com uma só categoria de entidade-NER

Data: 15/11/2023

Parâmetros diferentes:

i) Notebook: IDEA-ETAPA 3-FineTuneBERT_Spacy_NER.ipynb

Vocabulário: en_core_web_sm

Corpus: Materiais químicos (NER)

ii) Arquivo de configuração: config_spacy.cfg

[nlp]

lang = "en"

```
[components.transformer.model]
name = "allenai/scibert_scivocab_uncased"
```

```
[training]
dropout = 0.2
```

iii) Resultados

E	#	LOSS TRANS...	LOSS NER	ENTS_F	ENTS_P	ENTS_R	SCORE
0	0	833.55	573.11	2.89	2.39	3.66	0.03
25	200	53703.27	46078.65	77.94	76.70	79.21	0.78
50	400	613.66	654.79	78.61	77.77	79.47	0.79
75	600	147.71	149.65	79.81	77.24	82.55	0.80
100	800	95.52	91.93	78.44	75.87	81.19	0.78
125	1000	75.65	76.95	78.70	76.82	80.68	0.79
150	1200	97.55	87.84	79.15	77.04	81.39	0.79
175	1400	120.49	92.61	79.20	77.95	80.49	0.79
200	1600	70.02	63.45	78.61	75.33	82.19	0.79
225	1800	77.25	58.13	78.76	80.78	76.83	0.79
250	2000	33.40	29.00	79.60	77.23	82.12	0.80
275	2200	17.09	16.86	80.14	78.22	82.15	0.80
300	2400	26.81	22.13	80.21	78.25	82.27	0.80
325	2600	19.69	14.36	80.23	78.71	81.80	0.80
350	2800	26.12	29.63	80.33	80.52	80.13	0.80
375	3000	39.68	41.19	78.45	76.66	80.32	0.78
400	3200	63.92	50.65	79.27	77.55	81.07	0.79
425	3400	42.54	28.22	78.55	77.53	79.60	0.79
450	3600	84.52	62.57	78.55	77.23	79.92	0.79
475	3800	33.33	22.11	78.09	77.44	78.75	0.78
500	4000	52.39	38.86	77.40	77.89	76.92	0.77
525	4200	51.07	38.54	77.84	77.63	78.05	0.78
550	4400	20.15	11.72	78.70	76.97	80.52	0.79

```
-----
TOK    100.00
NER P  77.02
NER R  77.80
NER F  77.41
SPEED  3573
```

```
*****
*****
```

Corpus: -Corpus-2-Quimico-Com uma só categoria de entidade-RE

Data: 15/11/2023

Parâmetros diferentes:

i) Notebook: IDEA-ETAPA 4-FineTuneBERT_Spacy_RE.ipynb

Vocabulário: en_core_web_sm

Corpus: Materiais químicos (RE)

ii) Arquivo de configuração: \configs\config_spacy.cfg

Linha 12: lang = "en"

Linha 19: batch_size = 500

Linha 30: name = "roberta-base"

Linha 40: threshold = 0.5

Linha 60: max_length = 30

Linha 84: dropout = 0.2

Resultados do treinamento

E	#	LOSS TRANS...	LOSS RELAT...	REL_MICRO_P	REL_MICRO_R	REL_MICRO_F	SCORE
---	---	---------------	---------------	-------------	-------------	-------------	-------

0	0	2.93	1.92	0.62	76.70	1.24	0.01
12	100	115.73	40.13	0.00	0.00	0.00	0.00
25	200	0.21	3.30	41.81	26.04	32.09	0.32
37	300	0.33	2.40	51.92	24.00	32.83	0.33
50	400	0.12	1.51	41.37	41.68	41.53	0.42
62	500	0.05	1.15	42.58	39.79	41.14	0.41
75	600	0.03	1.01	44.44	38.68	41.36	0.41
87	700	0.02	0.96	47.98	35.98	41.12	0.41
100	800	0.03	0.99	36.55	50.41	42.38	0.42
112	900	0.03	0.99	44.22	43.01	43.61	0.44
125	1000	0.02	0.94	46.65	37.32	41.46	0.41

Random baseline:

threshold 0.00	{'rel_micro_p': '0.62', 'rel_micro_r': '100.00', 'rel_micro_f': '1.23'}
threshold 0.05	{'rel_micro_p': '0.62', 'rel_micro_r': '95.88', 'rel_micro_f': '1.24'}
threshold 0.10	{'rel_micro_p': '0.61', 'rel_micro_r': '89.38', 'rel_micro_f': '1.22'}
threshold 0.20	{'rel_micro_p': '0.61', 'rel_micro_r': '79.88', 'rel_micro_f': '1.22'}
threshold 0.30	{'rel_micro_p': '0.61', 'rel_micro_r': '69.50', 'rel_micro_f': '1.21'}
threshold 0.40	{'rel_micro_p': '0.60', 'rel_micro_r': '59.00', 'rel_micro_f': '1.19'}
threshold 0.50	{'rel_micro_p': '0.59', 'rel_micro_r': '48.25', 'rel_micro_f': '1.17'}
threshold 0.60	{'rel_micro_p': '0.59', 'rel_micro_r': '38.12', 'rel_micro_f': '1.15'}
threshold 0.70	{'rel_micro_p': '0.58', 'rel_micro_r': '28.25', 'rel_micro_f': '1.13'}
threshold 0.80	{'rel_micro_p': '0.59', 'rel_micro_r': '19.38', 'rel_micro_f': '1.14'}

threshold 0.90	{'rel_micro_p': '0.62', 'rel_micro_r': '10.25', 'rel_micro_f': '1.17'}
threshold 0.99	{'rel_micro_p': '0.76', 'rel_micro_r': '1.25', 'rel_micro_f': '0.95'}
threshold 1.00	{'rel_micro_p': '0.00', 'rel_micro_r': '0.00', 'rel_micro_f': '0.00'}

Results of the trained model:

threshold 0.00	{'rel_micro_p': '0.62', 'rel_micro_r': '100.00', 'rel_micro_f': '1.23'}
threshold 0.05	{'rel_micro_p': '29.98', 'rel_micro_r': '61.38', 'rel_micro_f': '40.28'}
threshold 0.10	{'rel_micro_p': '33.80', 'rel_micro_r': '56.75', 'rel_micro_f': '42.37'}
threshold 0.20	{'rel_micro_p': '37.89', 'rel_micro_r': '51.62', 'rel_micro_f': '43.70'}
threshold 0.30	{'rel_micro_p': '41.15', 'rel_micro_r': '47.38', 'rel_micro_f': '44.04'}
threshold 0.40	{'rel_micro_p': '42.06', 'rel_micro_r': '43.38', 'rel_micro_f': '42.71'}
threshold 0.50	{'rel_micro_p': '43.28', 'rel_micro_r': '40.62', 'rel_micro_f': '41.91'}
threshold 0.60	{'rel_micro_p': '45.44', 'rel_micro_r': '38.00', 'rel_micro_f': '41.39'}
threshold 0.70	{'rel_micro_p': '45.90', 'rel_micro_r': '34.25', 'rel_micro_f': '39.23'}
threshold 0.80	{'rel_micro_p': '47.59', 'rel_micro_r': '29.62', 'rel_micro_f': '36.52'}
threshold 0.90	{'rel_micro_p': '52.17', 'rel_micro_r': '24.00', 'rel_micro_f': '32.88'}
threshold 0.99	{'rel_micro_p': '54.39', 'rel_micro_r': '3.88', 'rel_micro_f': '7.23'}
threshold 1.00	{'rel_micro_p': '0.00', 'rel_micro_r': '0.00', 'rel_micro_f': '0.00'}

Corpus: -Corpus-1-SCIERC-Com "N" só categoria de entidade-NER

Data: 17/11/2023

Parâmetros diferentes:

i) Notebook: IDEA-ETAPA 3-FineTuneBERT_Spacy_NER.ipynb

Vocabulário: en_core_web_sm

Corpus: SCIERC

ii) Arquivo de configuração: config_spacy.cfg

[nlp]

Linha 12: lang = "en"

[components.transformer.model]

Linha 52: name = "allenai/scibert_scivocab_uncased"

[training]

dropout = 0.2

iii) Resultados

E	#	LOSS TRANS...	LOSS NER	ENTS_F	ENTS_P	ENTS_R	SCORE
---	----	-----	-----	-----	-----	-----	-----

0	0	362.95	254.67	0.28	0.17	0.71	0.00
28	200	551467.66	120390.53	28.12	31.87	25.16	0.28
57	400	70303.56	45843.45	62.87	64.01	61.76	0.63
85	600	442.70	30132.71	62.98	64.07	61.93	0.63
114	800	148.96	29685.84	63.57	64.74	62.44	0.64
142	1000	198.41	29678.24	63.21	63.27	63.15	0.63
171	1200	183.34	29111.92	63.11	61.66	64.62	0.63
200	1400	91.27	28892.72	61.83	62.29	61.38	0.62
228	1600	357.69	28714.81	62.56	62.59	62.53	0.63
257	1800	613.70	28449.70	61.56	59.64	63.60	0.62
285	2000	29284.82	32103.38	59.62	59.86	59.38	0.60
314	2200	6439.95	28732.28	62.23	61.97	62.49	0.62
342	2400	50.89	26979.76	62.65	62.62	62.67	0.63

TOK 100.00
NER P 65.11
NER R 63.20
NER F 64.14
SPEED 3441

===== NER (per type)
=====

	P	R	F
Method	71.95	70.10	71.01
Task	62.67	66.67	64.61
Generic	71.43	77.69	74.43
OtherScientificTerm	57.26	53.18	55.15
Material	61.42	49.06	54.55
Metric	66.67	57.14	61.54

Corpus: -Corpus-1-SCIERC-Com N categorias de entidade-RE
Data: 17/11/2023

Parâmetros diferentes:

i) Notebook: IDEA-ETAPA 3-FineTuneBERT_Spacy_RE.ipynb

Vocabulário: en_core_web_sm
Corpus: SCIERC (RE)

ii) Arquivo de configuração: \configs\config_spacy.cfg

Linha 12: lang = "en"

Linha 19: batch_size = 500
 Linha 30: name = "allenai/scibert_scivocab_uncased"
 Linha 40: threshold = 0.5
 Linha 60: max_length = 30
 Linha 84: dropout = 0.2

Resultados do treinamento

E # LOSS TRANS... LOSS RELAT... REL_MICRO_P REL_MICRO_R
 REL_MICRO_F SCORE

0	0	0.55	2.69	1.17	36.79	2.28	0.02
14	100	52.25	46.40	44.47	6.14	10.79	0.11
28	200	0.95	8.57	47.53	14.08	21.72	0.22
42	300	0.48	5.21	53.88	16.60	25.38	0.25
57	400	0.17	3.45	51.30	19.06	27.79	0.28
71	500	0.04	2.82	45.77	23.87	31.38	0.31
85	600	0.04	2.76	47.59	19.69	27.85	0.28
100	700	0.03	2.67	47.46	21.41	29.51	0.30
114	800	0.01	2.55	47.90	22.68	30.78	0.31
128	900	0.01	2.53	44.84	24.10	31.35	0.31
142	1000	0.01	2.50	46.63	24.10	31.78	0.32

Random baseline:

threshold 0.00	{'rel_micro_p': '1.57', 'rel_micro_r': '100.00', 'rel_micro_f': '3.09'}
threshold 0.05	{'rel_micro_p': '1.55', 'rel_micro_r': '94.09', 'rel_micro_f': '3.06'}
threshold 0.10	{'rel_micro_p': '1.56', 'rel_micro_r': '89.43', 'rel_micro_f': '3.07'}
threshold 0.20	{'rel_micro_p': '1.56', 'rel_micro_r': '79.90', 'rel_micro_f': '3.07'}
threshold 0.30	{'rel_micro_p': '1.54', 'rel_micro_r': '68.70', 'rel_micro_f': '3.01'}
threshold 0.40	{'rel_micro_p': '1.50', 'rel_micro_r': '57.20', 'rel_micro_f': '2.92'}
threshold 0.50	{'rel_micro_p': '1.50', 'rel_micro_r': '47.88', 'rel_micro_f': '2.92'}
threshold 0.60	{'rel_micro_p': '1.51', 'rel_micro_r': '38.13', 'rel_micro_f': '2.90'}
threshold 0.70	{'rel_micro_p': '1.52', 'rel_micro_r': '28.60', 'rel_micro_f': '2.88'}
threshold 0.80	{'rel_micro_p': '1.57', 'rel_micro_r': '19.69', 'rel_micro_f': '2.90'}
threshold 0.90	{'rel_micro_p': '1.65', 'rel_micro_r': '10.57', 'rel_micro_f': '2.86'}
threshold 0.99	{'rel_micro_p': '1.76', 'rel_micro_r': '1.14', 'rel_micro_f': '1.38'}
threshold 1.00	{'rel_micro_p': '1.72', 'rel_micro_r': '0.10', 'rel_micro_f': '0.20'}

Results of the trained model:

threshold 0.00	{'rel_micro_p': '1.57', 'rel_micro_r': '100.00', 'rel_micro_f': '3.09'}
threshold 0.05	{'rel_micro_p': '33.08', 'rel_micro_r': '35.75', 'rel_micro_f': '34.36'}
threshold 0.10	{'rel_micro_p': '36.52', 'rel_micro_r': '31.30', 'rel_micro_f': '33.71'}
threshold 0.20	{'rel_micro_p': '40.40', 'rel_micro_r': '27.46', 'rel_micro_f': '32.70'}
threshold 0.30	{'rel_micro_p': '44.14', 'rel_micro_r': '25.39', 'rel_micro_f': '32.24'}
threshold 0.40	{'rel_micro_p': '46.89', 'rel_micro_r': '23.42', 'rel_micro_f': '31.24'}
threshold 0.50	{'rel_micro_p': '48.27', 'rel_micro_r': '20.21', 'rel_micro_f': '28.49'}

threshold 0.60	{'rel_micro_p': '48.48', 'rel_micro_r': '18.13', 'rel_micro_f': '26.40'}
threshold 0.70	{'rel_micro_p': '50.96', 'rel_micro_r': '16.58', 'rel_micro_f': '25.02'}
threshold 0.80	{'rel_micro_p': '53.73', 'rel_micro_r': '14.20', 'rel_micro_f': '22.46'}
threshold 0.90	{'rel_micro_p': '54.40', 'rel_micro_r': '10.26', 'rel_micro_f': '17.26'}
threshold 0.99	{'rel_micro_p': '54.24', 'rel_micro_r': '3.32', 'rel_micro_f': '6.25'}
threshold 1.00	{'rel_micro_p': '62.50', 'rel_micro_r': '1.04', 'rel_micro_f': '2.04'}
