TFM - Cuaderno de Bitácora

Germán García García

https://github.com/3Gsman/DeepTerminologyExtraction

Temas para la reunión

Datasets

https://github.com/LIAAD/KeywordExtractor-Datasets

INSPEC:
Archivos: 2000

Abstracts de artículos de revistas de informática entre 1998 y 2002.
Cada documento tiene 2 conjuntos de keywords:

- Controlled keywords: manualmente controladas que aparecen en el tesauro de Inspec pero que pueden no aparecer en el documento.
- Uncontrolled keywords: Asignadas libremente por los editores (no se limitan al tesauro o al documento)

KDD:
Archivos: 755

Abstracts de papers de la conferencia Knowledge Discovery and Data Mining, publicados entre 2004 y 2014. Los keywords de los papers son etiquetados por el autor.

WWW:
Archivos: 1330

Abstracts de la conferencia World Wide Web, entre 2004 y 2014, Los keywords de los papers son etiquetados por el autor.

SemEval2010:
Archivos: 244

Artículos completos de ACM Digital Library. Entre 6 y 8 páginas sobre 4 campos de la informática. Cada artículo tiene un conjunto de palabras clave asignadas por el autor y un conjunto de palabras clave asignadas por editores profesionales, las cuales pueden aparecer o no explícitamente en el texto.

SemEval2017:
Archivos: 500

Párrafos seleccionados de revistas con artículos, sobre informática, física y ciencias. Las keywords son seleccionadas por un experto y un estudiante sin graduar no experto.

Kp20k
Archivos: 568000

Abstracts de artículos académicos con 528000 training, 20000 validación y 20000 test.

Tabla datasets:

Dataset
Número de palabras abstract (max, min, media)
Número de palabras clave
Número de palabras en las palabras clave
Bi grams, tri grams

```
INSPEC
Abstracts number: 2000
Keys number: 2000
Words per abstract
Media: 124.364
STD: 59.13150405144192
Min: 15
Max: 502
Keywords per abstract
Media: 14.11
STD: 6.408153573074545
Min: 2
Max: 40
Words per keywords
Media: 2.224060949681077
STD: 0.8596185587760682
Min: 1
Max: 9

SEMEVAL2010
Abstracts number: 243
Keys number: 243
Words per abstract
Media: 8154.6543209876545
STD: 1973.5407779740183
Min: 2508
Max: 14471
Keywords per abstract
Media: 15.576131687242798
STD: 3.7446057973608027
Min: 8
Max: 38
Words per keywords
Media: 2.158784676354029
STD: 0.8897422260009823
Min: 1
Max: 9

KDD
```

```
Abstracts number: 755
Keys number: 755
Words per abstract
Media: 74.0794701986755
STD: 84.06142152365945
Min: 7
Max: 329
Keywords per abstract
Media: 4.096688741721854
STD: 1.7335592715039714
Min: 1
Max: 13
Words per keywords
Media: 1.978338182993857
STD: 0.7829746943623613
Min: 1
Max: 8

WWW
Abstracts number: 1330
Keys number: 1330
Words per abstract
Media: 82.03834586466165
STD: 80.4972462566211
Min: 6
Max: 346
Keywords per abstract
Media: 4.815789473684211
STD: 2.0850733298519186
Min: 1
Max: 21
Words per keywords
Media: 1.9017954722872756
STD: 0.8578939896338169
Min: 1
Max: 8
```

kp20k

```
570809
570809
Words per abstract
Media: 147.74611647678995
STD: 67.99700836029831
Min: 0
Max: 2996
Keywords per abstract
Media: 5.286596742518075
STD: 3.774151174068877
Min: 1
```

```
Max: 110
Words per keywords
Media: 1.9438113331722802
STD: 0.9490471084333197
Min: 0
Max: 9
```

Conjunto de datos		Inspec	SemEval2010	KDD	WWW	kp20k
Número de artículos		2000	243	755	1330	570809
Palabras por artículos	Media	124,364	8154,65	74,07	82,03	147,74
	Desviación	59,13	1973,54	84,06	80,49	67,99
	Min	15	2508	7	6	0
	Max	502	14471	329	346	2996
Palabras clave por articulo	Media	14,11	15,57	4,09	4,81	5,28
	Desviación	6,4	3,74	1,73	2,08	3,77
	Min	2	8	1	1	1
	Max	40	38	13	21	110
Palabras por palabras clave	Media	2,22	2,15	1,97	1,9	1,94
	Desviación	0,85	0,88	0,78	0,85	0,94
	Min	1	1	1	1	1
	Max	9	9	8	8	92

Tablas y gráfica estado del arte

INSPEC

Modelo	Año	Pre	Rec	F1	Comentarios
n-gram w. tag [1]	2003	0.252	0.517	0.339	Utiliza conocimiento lingüístico para la representación, frente el uso único de estadísticas (como puede ser term frequency). ML
SC [2]	2010	0.350	0.660	0.457	Sin supervisar. Spectral Clustering Se agrupan términos ejemplares según la relación entre los términos y se utilizan los términos ejemplares para extraer las keyph.
TopicRank [3]	2013	0.276	0.315	0.279	Basado en grafos. Clusters de temas usados como vértices en un grafo. Se le asigna una puntuación de importancia a cada tema. Las keyphr se seleccionan dependiendo de los temas con más importancia.
SGRank [4]	2015	0.27	0.46	0.3395	Híbrido estadístico - gráfico. Se realizan n-grams, se puntúan según una versión de tf-idf, los mejores candidatos se analizan según otras heurísticas estadísticas (posición de la primera ocurrencia, longitud del término) se le aplica un algoritmo basado en grafo que devuelve el rank. **Valores Pre y Rec aprox*
GloVe-100 [6]	2018	0.34	0.578	0.428	Bidirectional Long Short-Term Memory Recurrent Neural Network. Usa GloVe Embeddings, Interesante como se explica.
Key2Vec [7]	2018	0.5758	0.4209	0.4863	Utiliza embeddings específicos del dominio para mejorar los resultados. Utiliza un método basado en grafos parecido a textrank.
HATR [14]	2019	-	-	0.3005	BERT + BiLSTM-CRF + document-level attention (extrae la información más relevante de cada oración)
BiLSTM-CRF [8]	2019	-	-	0.593	Utiliza BiLSTM-CRF + contextualized embeddings (SciBERT)
YAKE [13]	2020	-	-	0.316	Hay pseudo código y se realizan muchas pruebas en distintos datasets. Sistema no supervisado que utiliza métricas estadísticas de las palabras.
SKE-Base-Cls [16]	2020	-	-	0.342	span-based feature representation. Utiliza BERT y 2 Bi-LSTM.

Ref	Autor	Año	Prec	Exha	F1
SC [8]	Liu Z	2010	0.350	0.660	0.457
TopicRank [25]	Bougouin A	2013	0.276	0.315	0.279
SGRank [22]	Danesh S	2015	0.27	0.46	0.3395
GloVe-100 [9]	Basaldella M	2018	0.34	0.578	0.428
Key2Vec [41]	Mahata D	2018	0.5758	0.4209	0.4863
HATR [42]	Yingyi Zhang	2019	-	-	0.3005
BiLSTM-CRF [35]	Sahrawat D	2019	-	-	0.593
YAKE [40]	Campos R	2020	-	-	0.316
SKE-Base-Cls [39]	Mu F	2020	-	-	0.342

SemEval2010

Modelo	Año	Pre	Rec	F1	Comentarios
HUMB [9]	2010	0.248	0.206	0.225	Utiliza varios modelos de ML (Decision tree, Multi-Layer perceptron, Support Vector Machine)
TopicRank [3]	2013	0.149	0.103	0.121	Basado en grafos. Clusters de temas usados como vértices en un grafo. Se le asigna una puntuación de importancia a cada tema. Las keyphr se seleccionan dependiendo de los temas con más importancia.
SGRank [4]	2015	0.26	0.26	0.2607	Híbrido estadístico - gráfico. Se realizan n-grams, se puntúan según una versión de tf-idf, los mejores candidatos se analizan según otras heurísticas estadísticas (posición de la primera ocurrencia, longitud del término) se le aplica un algoritmo basado en grafo que devuelve el rank. **Valores Pre y Rec aprox*
Key2Vec [7]	2018	0.3529	0.2467	0.2904	Utiliza embeddings específicos del dominio para mejorar los resultados. Utiliza un método basado en grafos parecido a textrank.
BiLSTM-CRF [8]	2019	-	-	0.357	Utiliza BiLSTM-CRF + contextualized embeddings (SciBERT)
YAKE [13]	2020	-	-	0.211	Hay pseudo código y se realizan muchas pruebas en distintos datasets. Sistema no supervisado que utiliza métricas estadísticas de las palabras.
SKE-Large-CIs [16]	2020	-	-	0.358	span-based feature representation. Utiliza BERT y 2 Bi-LSTM.

Ref	Autor	Año	Prec	Exha	F1
HUMB [11]	Lopez P	2010	0.248	0.206	0.225
TopicRank [25]	Bougouin A	2013	0.149	0.103	0.121
SGRank [22]	Danesh S	2015	0.26	0.26	0.26
Key2Vec [41]	Mahata D	2018	0.3529	0.2467	0.2904
BiLSTM-CRF [35]	Sahrawat D	2019	-	-	0.357
YAKE [40]	Campos R	2020	-	-	0.211
SKE-Large-CIs [39]	Mu F	2020	-	-	0.358

WWW

Modelo	Año	Pre	Rec	F1	Comentarios
CRF [10]	2016	0.2669	0.2067	0.2112	Uso de una capa CRF. Método de sequence modeling. No pone el uso de 10 keyphrases
MIKE [11]	2017	0.0907	0.1814	0.1209	Modelo gráfico. Resuelve el problema a modo de problema de optimización..
Bi-LSTM-CRF [12]	2019	0.6433	0.2843	0.3943	Metodo squence labeling. Utiliza Bi-LSTM-CRF. Muestra varios esquemas de la infraestructura. Y bien explicado sus elementos. Utiliza 100-dimension Glove
YAKE	2020	-	-	0.172	Hay pseudo código y se realizan muchas pruebas en distintos datasets. Sistema no supervisado que utiliza

[13]					métricas estadísticas de las palabras.
WEKE(WE t) [17]	2020	0.1476	0.15	0.1488	p = 5 // Nuevo modelo de word embeddings, que captura información local de contexto (patrón entre palabras en el nivel semántico).

Ref	Autor	Año	Prec	Exha	F1
CRF [43]	Das Gollapalli S	2016	0.2669	0.2067	0.2112
MIKE [26]	Yuxiang Zhang	2017	0.0907	0.1814	0.1209
Bi-LSTM-CRF [34]	Rabah A	2019	0.6433	0.2843	0.3943
WEKE(WE t) [44]	Yuxiang Zhang	2020	0.1476	0.15	0.1488
YAKE [40]	Campos R	2020	-	-	0.172

KDD

Modelo	Año	Pre	Rec	F1	Comentarios
CRF [10]	2016	0.2933	0.2343	0.2417	Uso de una capa CRF. Método de sequence modeling.No pone el uso de 10 keyphrases
MIKE [11]	2017	0.0933	0.2242	0.1318	Modelo gráfico. Resuelve el problema a modo de problema de optimización..
Bi-LSTM-CRF [12]	2019	0.5783	0.3185	0.4108	Método sequence labeling. Utiliza Bi-LSTM-CRF. Muestra varios esquemas de la infraestructura. Y bien explicado sus elementos. Utiliza 100-dimension Glove
YAKE [13]	2020	-	-	0.156	Hay pseudo código y se realizan muchas pruebas en distintos datasets. Sistema no supervisado que utiliza métricas estadísticas de las palabras.
WEKE(WE t) [17]	2020	0.1588	0.157	0.1579	p = 4 // Nuevo modelo de word embeddings, que captura información local de contexto (patrón entre palabras en el nivel semántico).

Ref	Autor	Año	Prec	Exha	F1
CRF [43]	Das Gollapalli S	2016	0.2933	0.2343	0.2417
MIKE [26]	Yuxiang Zhang	2017	0.0933	0.2242	0.1318
Bi-LSTM-CRF [34]	Rabah A	2019	0.5783	0.3185	0.4108
WEKE(WE t) [44]	Yuxiang Zhang	2020	0.1588	0.157	0.1579
YAKE [40]	Campos R	2020	-	-	0.156

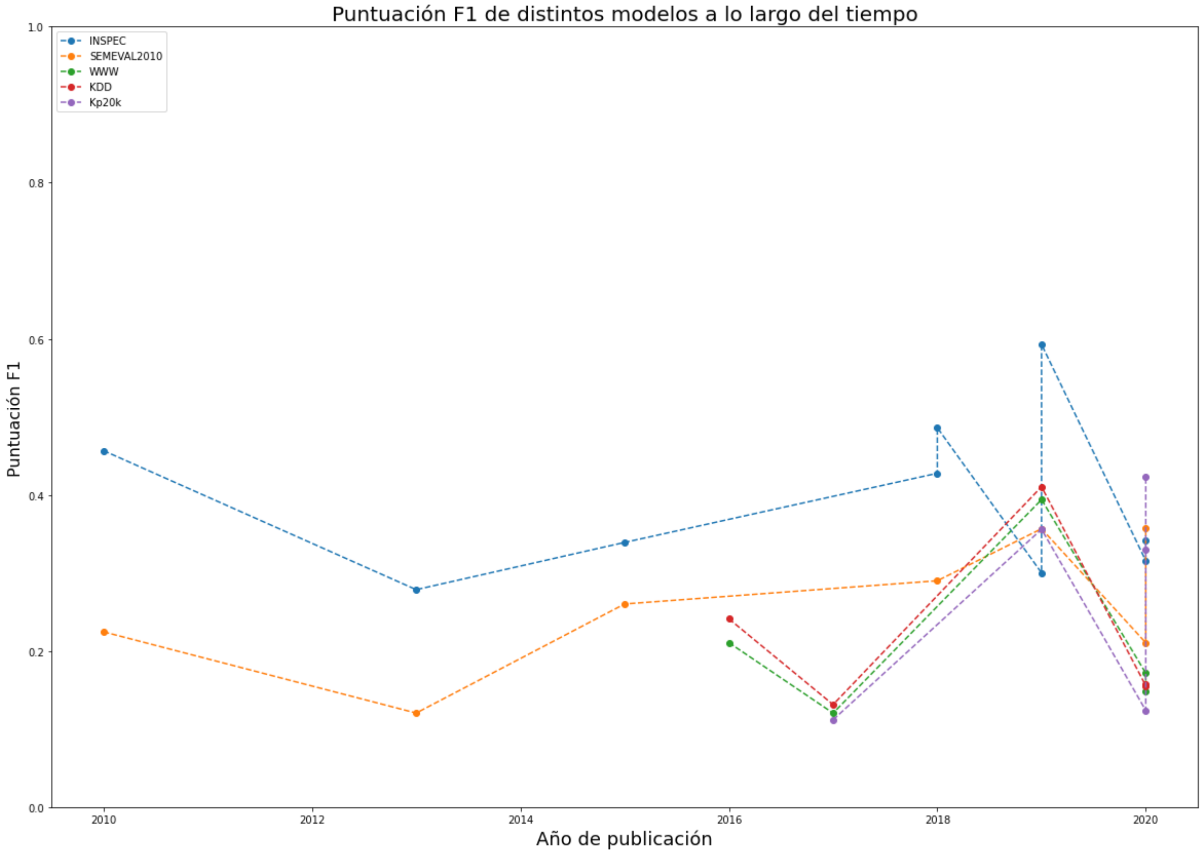
Kp20k

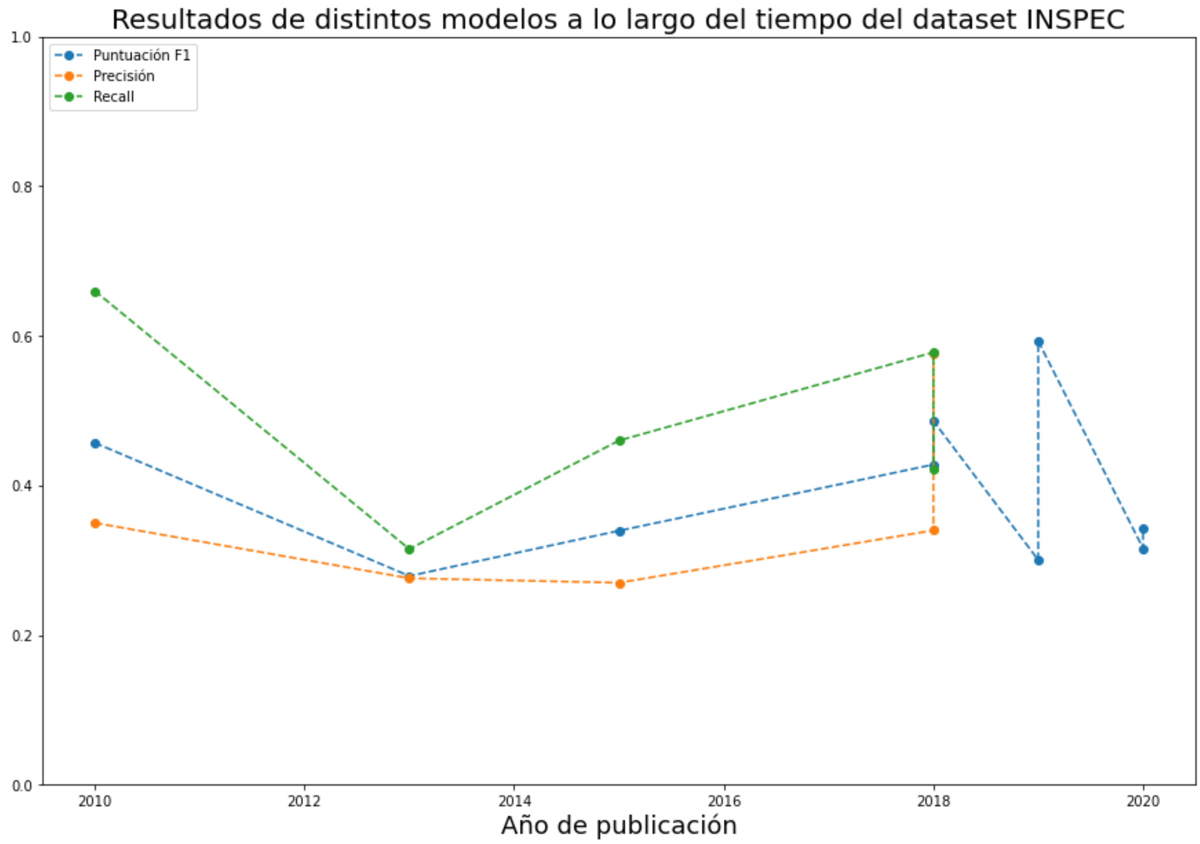
Modelo	Año	Pre	Rec	F1	Comentarios
HATR [14]	2017	-	-	0.1124	BERT + BiLSTM-CRF + document-level attention (extrae la información más relevante de cada oración)
Bi-LSTM-CRF [12]	2019	0.6419	0.2466	0.3563	Metodo squence labeling. Utiliza Bi-LSTM-CRF. Muestra varios esquemas de la infraestructura. Y bien explicado sus elementos. Utiliza 100-dimension Glove
WEKE(WE	2020	0.1167	0.1329	0.1243	p = 6 // Nuevo modelo de word embeddings, que captura

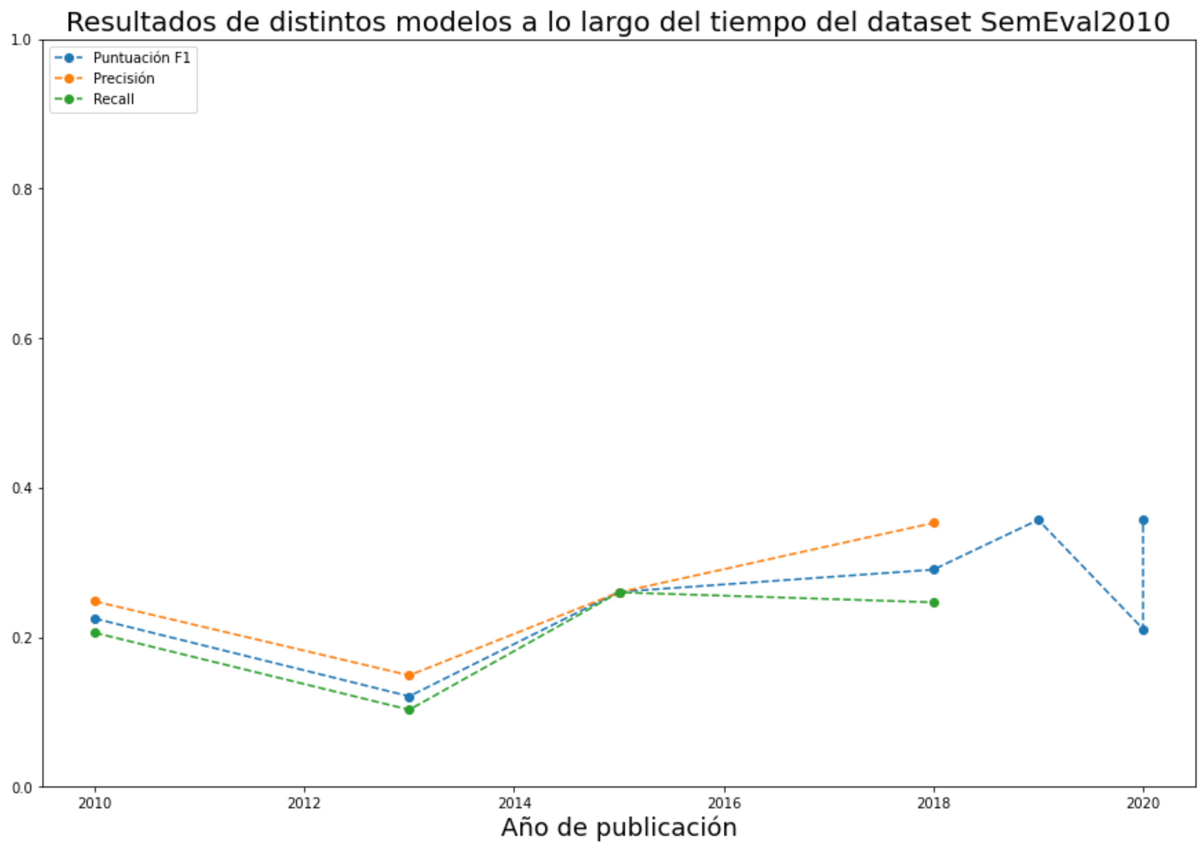
t) [17]					información local de contexto (patrón entre palabras en el nivel semántico).
SKE-Large -CIs [16]	2020	-	-	0.330	span-based feature representation. Utiliza BERT y 2 Bi-LSTM.
DAKE [15]	2020	0.6821	0.3066	0.423	Utiliza BERT y BiLSTM-CRF. Bien explicado cada capa y fase. Slides.

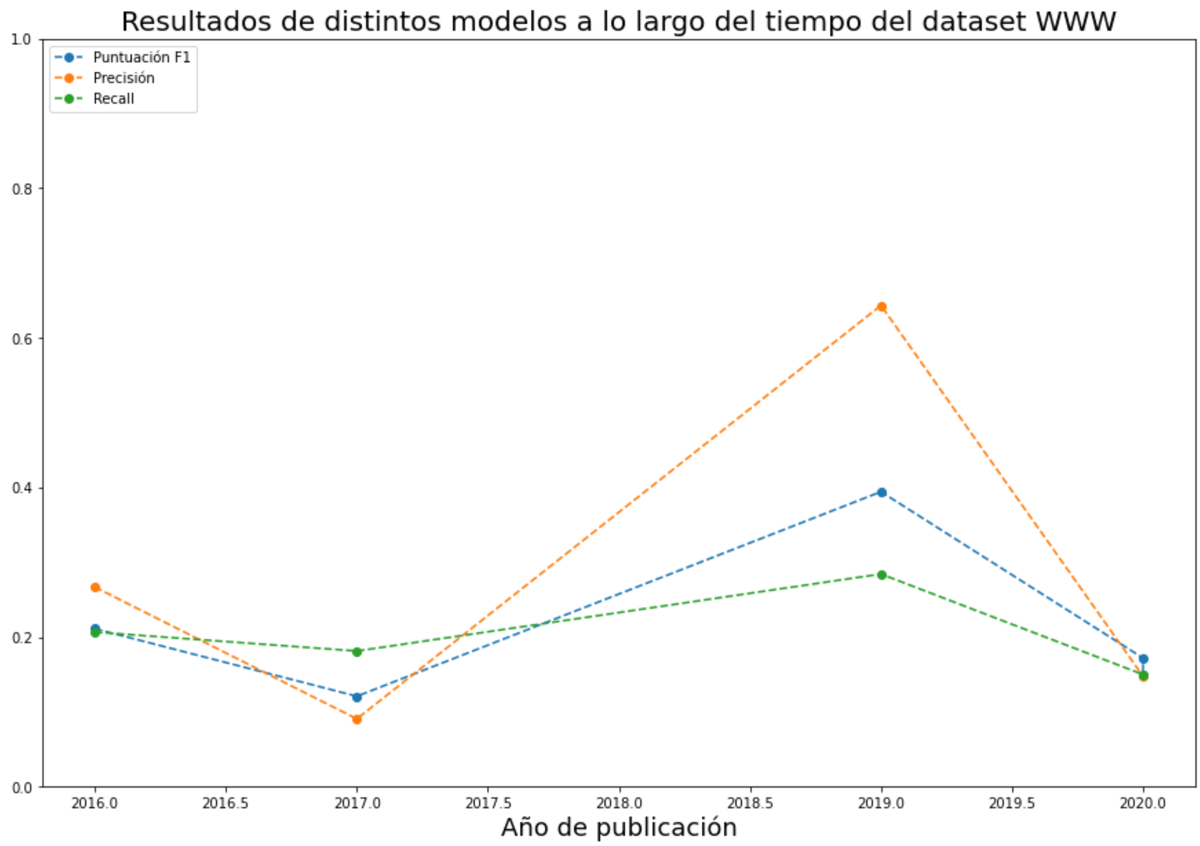
Ref	Autor	Año	Prec	Exha	F1
Bi-LSTM-CRF [34]	Rabah A	2019	0.6419	0.2466	0.3563
WEKE(WE t) [44]	Yuxiang Zhang	2020	0.1167	0.1329	0.1243
SKE-Large-CIs [39]	Mu F	2020	-	-	0.330
DAKE [36]	Santosh T	2020	0.6821	0.3066	0.423

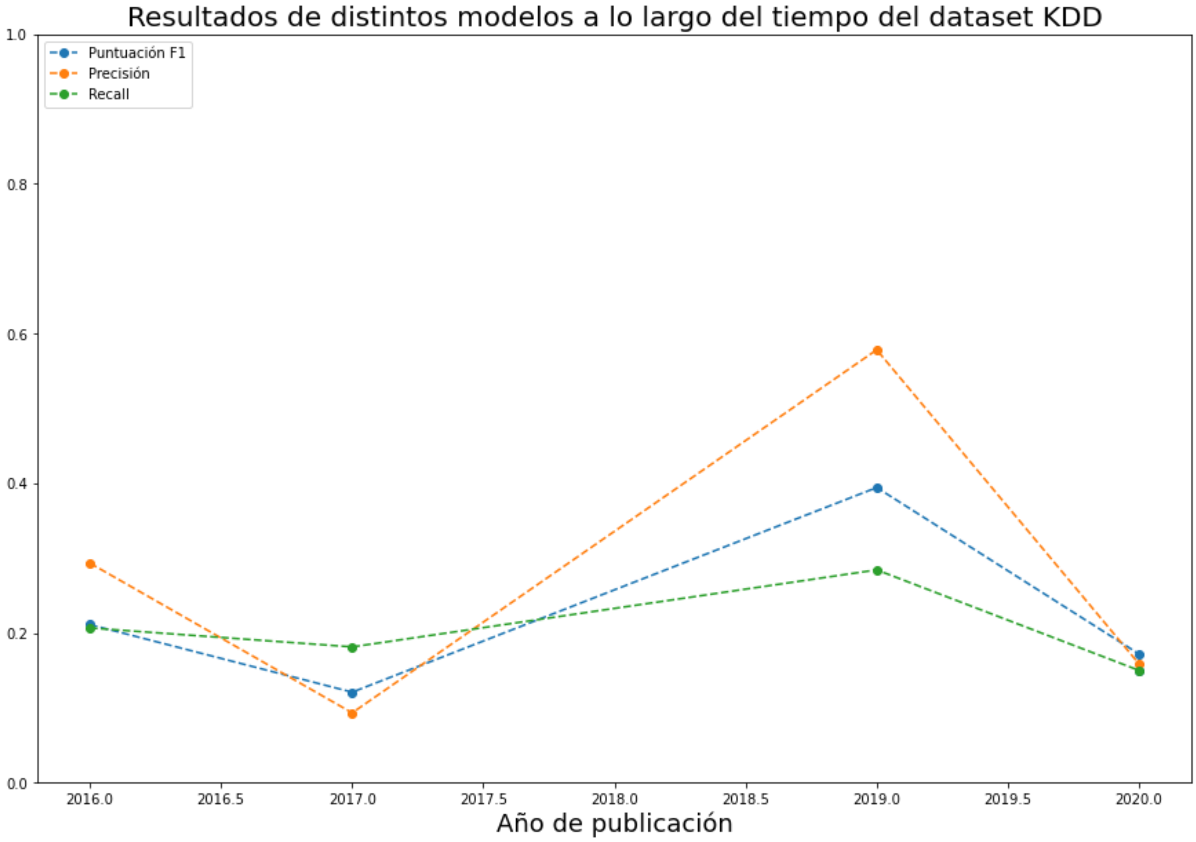
Gráficas:

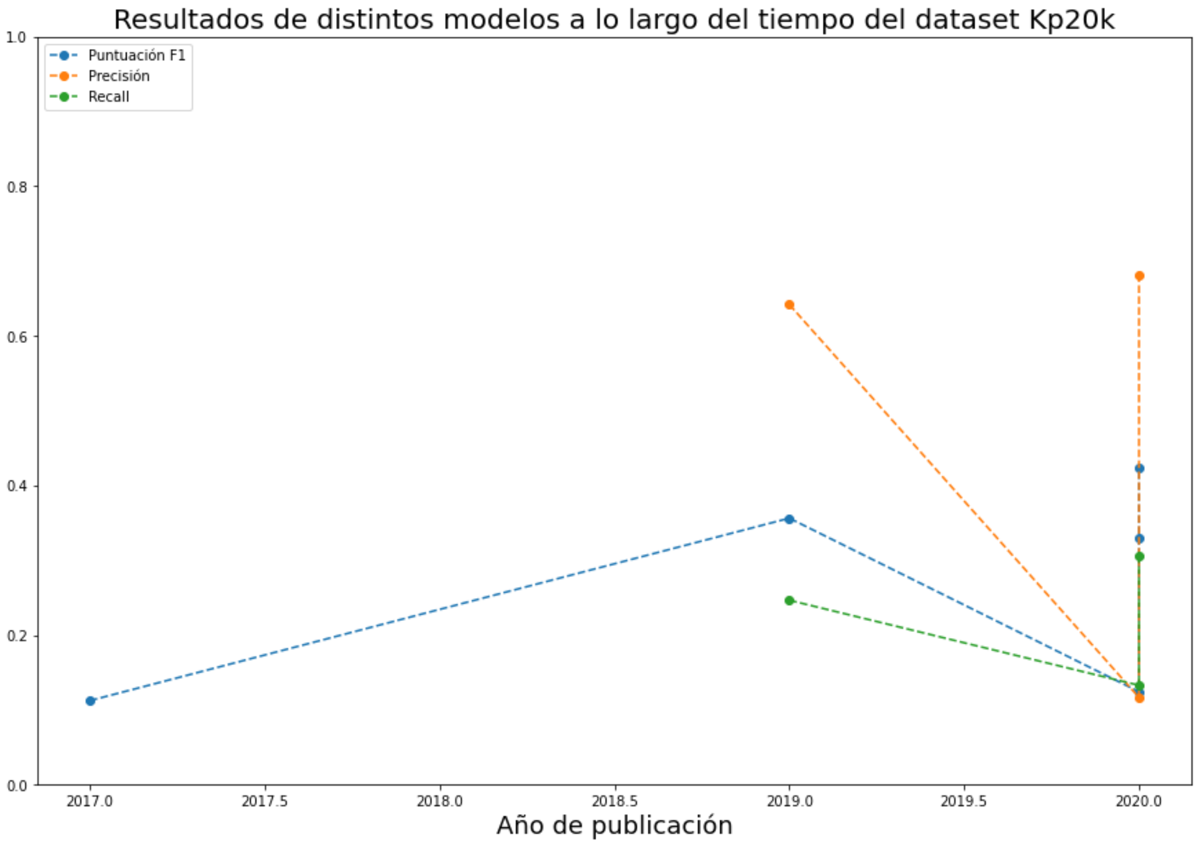

Puntuación F1 de distintos modelos a lo largo del tiempo
INSPEC
SEMEVAL2010
WWW
KDD
Kp20k
Puntuación F1
Año de publicación

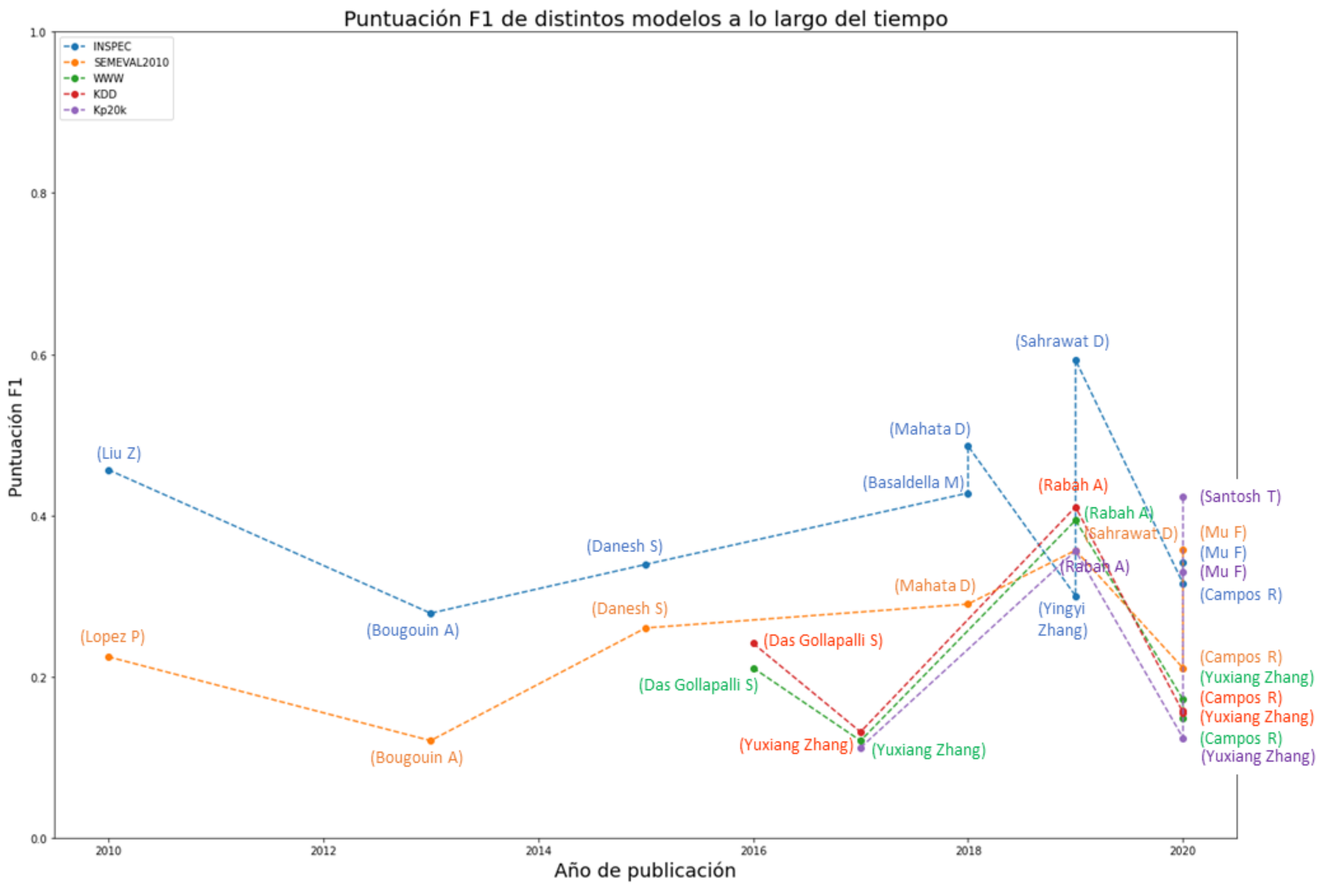

Resultados de distintos modelos a lo largo del tiempo del dataset INSPEC
Puntuación F1
Precisión
Recall
Año de publicación


Resultados de distintos modelos a lo largo del tiempo del dataset SemEval2010
Puntuación F1
Precisión
Recall
1.0
0.8
0.6
0.4
0.2
0.0
2010
2012
2014
2016
2018
2020
Año de publicación


Resultados de distintos modelos a lo largo del tiempo del dataset WWW
Puntuación F1
Precisión
Recall
1.0
0.8
0.6
0.4
0.2
0.0
2016.0
2016.5
2017.0
2017.5
2018.0
2018.5
2019.0
2019.5
2020.0
Año de publicación


Resultados de distintos modelos a lo largo del tiempo del dataset KDD
Puntuación F1
Precisión
Recall
1.0
0.8
0.6
0.4
0.2
0.0
2016.0
2016.5
2017.0
2017.5
2018.0
2018.5
2019.0
2019.5
2020.0
Año de publicación


Resultados de distintos modelos a lo largo del tiempo del dataset Kp20k
Puntuación F1
Precisión
Recall
1.0
0.8
0.6
0.4
0.2
0.0
2017.0
2017.5
2018.0
2018.5
2019.0
2019.5
2020.0
Año de publicación



Artículos:

[1] *Improved Automatic Keyword Extraction Given More Linguistic Knowledge* - Anette Hulth (2003)

[2] *Clustering to Find Exemplar Terms for Keyphrase Extraction* - Zhiyuan Liu, Peng Li, Yabin Zheng, Maosong Sun (2010)

[3] *TopicRank: Graph-Based Topic Ranking for Keyphrase Extraction* - Adrien Bougouin, Florian Boudin and Beatrice Daille (2013)

[4] *SGRank: Combining Statistical and Graphical Methods to Improve the State of the Art in Unsupervised Keyphrase Extraction* - Soheil Danesh, Tamara Sumner, James H. Martin (2015)

[6] *Bidirectional LSTM Recurrent Neural Network for Keyphrase Extraction* - Marco Basaldella, Elisa Antolli, Giuseppe Serra and Carlo Tasso (2018)

[7] *Key2Vec: Automatic Ranked Keyphrase Extraction from Scientific Articles using Phrase Embeddings* - Debanjan Mahata, John Kuriakose, Rajiv Ratn Shah, Roger Zimmermann (2018)

[8] *Keyphrase Extraction from Scholarly Articles as Sequence Labeling using Contextualized Embeddings* - Dhruva Sahrawat3, Debanjan Mahata1y, Raymond Zhang1 (2019)

[9] *HUMB: Automatic key term extraction from scientific articles in GROBID* - Lopez PRomary L (2010)

[10] *Keyphrase Extraction using Sequential Labeling* - Sujatha Das Gollapalli and Xiao-li Li (2016)

[11] *MIKE: Keyphrase Extraction by Integrating Multidimensional Information* - Yuxiang Zhang (2017)

[12] *Bi-LSTM-CRF Sequence Labeling for Keyphrase Extraction from Scholarly Documents* - Rabah A. Al-Zaidy, Cornelia Caragea, C. Lee Giles (2019)

[13] *YAKE! Keyword Extraction from Single Documents using Multiple Local Features. Information Sciences*. - Campos, R., Mangaravite, V., Pasquali, A., Jorge, A., Nunes, C., & Jatowt, A. (2020)

[14] *Unsupervised Keyphrase Extraction in Academic Publications Using Human Attention* - Yingyi Zhang and Chengzhi Zhang (2019)

[15] *DAKE: Document-Level Attention for Keyphrase Extraction* - Tokala Yaswanth Sri Sai Santosh1, Debarshi Kumar Sanyal2, Plaban Kumar Bhowmick3, and Partha Pratim Das1 (2020)

[16] *Keyphrase Extraction with Span-based Feature Representations* - Funan Mu], Zhenting Yu], LiFeng Wang (2020)

[17] *WEKE: Learning Word Embeddings for Keyphrase Extraction* - Yuxiang Zhang Email authorHuan LiuBei ShiXiaoli LiSuge Wang (2020)

Resultados

Prueba	Epochs	F1 training	F1 dev	F1 train	Tiempo	Comentarios
Prueba1	62/100	0.5892	0.5583	0.6980		
Roberta SharInsp	74/100	0.5762	0.5542	0.5941	91	
Bert SharInsp	65/100	0.5839	0.5609	0.7576	85	
Bert GermInsp	70/100	0.5552	0.5376	0.7776	157	
Bert Kp20k small	68/100	0.2121	0.2176	0.6343	160	

Raw:

INSPEC

Prueba 1
12:08 / 13:18

EPOCH 52 done: loss 15.7013 - lr 0.0001953
2021-05-28 13:18:22,492 DEV : loss 19.05150604248047 - score 0.5451

TRAINING
KEY tp: 2862 - fp: 2107 - fn: 1982 - precision: 0.5760 - recall: 0.5908 - f1-score: 0.5833

DEV
tp: 2797 - fp: 2519 - fn: 1769 - precision: 0.5261 - recall: 0.6126 - f1-score: 0.5661
tensor(19.1388, device='cuda:0')

TRAIN
tp: 7021 - fp: 4268 - fn: 2757 - precision: 0.6219 - recall: 0.7180 - f1-score: 0.6665
tensor(15.8236, device='cuda:0')

Prueba 2
12:09 / 13:29

EPOCH 62 done: loss 13.9009 - lr 0.0001953
2021-05-28 13:29:24,275 DEV : loss 19.519670486450195 - score 0.5424

TRAINING
tp: 2858 - fp: 2073 - fn: 1986 - precision: 0.5796 - recall: 0.5900 - f1-score: 0.5848

DEV
tp: 2627 - fp: 2175 - fn: 1939 - precision: 0.5471 - recall: 0.5753 - f1-score: 0.5608

TRAIN
tp: 6996 - fp: 3157 - fn: 2782 - precision: 0.6891 - recall: 0.7155 - f1-score: 0.7020
tensor(13.8435, device='cuda:0')

Prueba 3
12:09 / 13:17

EPOCH 51 done: loss 16.0274 - lr 0.0001953
2021-05-28 13:16:35,804 DEV : loss 18.840885162353516 - score 0.5527

TRAINING
tp: 3068 - fp: 2395 - fn: 1776 - precision: 0.5616 - recall: 0.6334 - f1-score: 0.5953

DEV
tp: 2797 - fp: 2519 - fn: 1769 - precision: 0.5261 - recall: 0.6126 - f1-score: 0.5661
tensor(19.1388, device='cuda:0')

TRAIN
tp: 7021 - fp: 4268 - fn: 2757 - precision: 0.6219 - recall: 0.7180 - f1-score: 0.6665
tensor(15.8236, device='cuda:0')

Prueba 4
12:08 / 13:18

EPOCH 53 done: loss 15.7839 - lr 0.0001953
2021-05-28 13:17:56,315 DEV : loss 18.922517776489258 - score 0.5509

TRAINING
tp: 3250 - fp: 2791 - fn: 1594 - precision: 0.5380 - recall: 0.6709 - f1-score: 0.5972

DEV
tp: 2951 - fp: 2974 - fn: 1615 - precision: 0.4981 - recall: 0.6463 - f1-score: 0.5626
tensor(19.8108, device='cuda:0')

TRAIN
tp: 7118 - fp: 5443 - fn: 2660 - precision: 0.5667 - recall: 0.7280 - f1-score: 0.6373
tensor(17.8116, device='cuda:0')

Prueba 5
12:08 / 13:22

EPOCH 56 done: loss 15.1926 - lr 0.0001953
2021-05-28 13:21:43,569 DEV : loss 18.910654067993164 - score 0.5543

TRAINING
tp: 2911 - fp: 2161 - fn: 1933 - precision: 0.5739 - recall: 0.6009 - f1-score: 0.5871

DEV
tp: 2633 - fp: 2222 - fn: 1933 - precision: 0.5423 - recall: 0.5767 - f1-score: 0.5590
tensor(18.7270, device='cuda:0')

TRAIN
tp: 6912 - fp: 3523 - fn: 2866 - precision: 0.6624 - recall: 0.7069 - f1-score: 0.6839
tensor(14.5071, device='cuda:0')

Prueba 6
07:47 / 09:03
EPOCH 58 done: loss 15.4786 - lr 0.0001953
2021-05-31 09:03:08,314 DEV : loss 19.03093910217285 - score 0.5444

TRAINING
KEY tp: 2849 - fp: 2079 - fn: 1995 - precision: 0.5781 - recall: 0.5882 - f1-score: 0.5831

DEV
tp: 2579 - fp: 2168 - fn: 1987 - precision: 0.5433 - recall: 0.5648 - f1-score: 0.5538
tensor(19.0716, device='cuda:0')

TRAIN
tp: 7320 - fp: 2947 - fn: 2458 - precision: 0.7130 - recall: 0.7486 - f1-score: 0.7304
tensor(12.5892, device='cuda:0')

Prueba 7
07:47 / 08:58
EPOCH 57 done: loss 14.5998 - lr 0.0001953
2021-05-31 08:58:12,684 DEV : loss 19.199932098388672 - score 0.5516

TRAINING
KEY tp: 2970 - fp: 2155 - fn: 1874 - precision: 0.5795 - recall: 0.6131 - f1-score: 0.5958

DEV
tp: 2655 - fp: 2276 - fn: 1911 - precision: 0.5384 - recall: 0.5815 - f1-score: 0.5591
tensor(18.8507, device='cuda:0')

TRAIN
tp: 6729 - fp: 3865 - fn: 3049 - precision: 0.6352 - recall: 0.6882 - f1-score: 0.6606
tensor(15.5457, device='cuda:0')

Prueba 8
07:47 / 09:06
EPOCH 61 done: loss 14.1091 - lr 0.0001953
2021-05-31 09:06:24,731 DEV : loss 19.550012588500977 - score 0.5399

TRAINING
KEY tp: 3009 - fp: 2280 - fn: 1835 - precision: 0.5689 - recall: 0.6212 - f1-score: 0.5939

DEV
tp: 2744 - fp: 2379 - fn: 1822 - precision: 0.5356 - recall: 0.6010 - f1-score: 0.5664
tensor(19.0746, device='cuda:0')

TRAIN
tp: 7120 - fp: 3826 - fn: 2658 - precision: 0.6505 - recall: 0.7282 - f1-score: 0.6871
tensor(14.6029, device='cuda:0')

Prueba 9
07:47 / 10:23
EPOCH 69 done: loss 12.4409 - lr 0.0001953
2021-05-31 10:22:22,620 DEV : loss 20.40585708618164 - score 0.5401

TRAINING
KEY tp: 2958 - fp: 2264 - fn: 1886 - precision: 0.5664 - recall: 0.6107 - f1-score: 0.5877

DEV
tp: 2696 - fp: 2369 - fn: 1870 - precision: 0.5323 - recall: 0.5905 - f1-score: 0.5599
tensor(19.3056, device='cuda:0')

TRAIN
tp: 7220 - fp: 3582 - fn: 2558 - precision: 0.6684 - recall: 0.7384 - f1-score: 0.7017
tensor(13.9425, device='cuda:0')

Prueba 10

07:47 / 10:09
EPOCH 65 done: loss 13.4039 - lr 0.0001953
2021-05-31 10:09:27,370 DEV : loss 19.732463836669922 - score 0.5474

TRAINING
KEY tp: 3043 - fp: 2396 - fn: 1801 - precision: 0.5595 - recall: 0.6282 - f1-score: 0.5919

DEV
tp: 2740 - fp: 2451 - fn: 1826 - precision: 0.5278 - recall: 0.6001 - f1-score: 0.5616
tensor(19.4120, device='cuda:0')

TRAIN
tp: 7412 - fp: 3666 - fn: 2366 - precision: 0.6691 - recall: 0.7580 - f1-score: 0.7108
tensor(13.5579, device='cuda:0')

6:56 / 8:09
EPOCH 56 done: loss 14.6564 - lr 0.0001953
2021-06-28 08:09:11,526 DEV : loss 19.129915237426758 - score 0.5482

KEY tp: 3041 - fp: 2377 - fn: 1803 - precision: 0.5613 - recall: 0.6278 - f1-score: 0.5927

KEY tp: 2709 - fp: 2479 - fn: 1857 - precision: 0.5222 - recall: 0.5933 - f1-score: 0.5555
tensor(19.1536, device='cuda:0')

KEY tp: 6954 - fp: 4205 - fn: 2824 - precision: 0.6232 - recall: 0.7112 - f1-score: 0.6643
tensor(15.9225, device='cuda:0')

6:51 / 8:22
EPOCH 65 done: loss 14.5846 - lr 0.0001953
2021-06-28 08:22:10,884 DEV : loss 19.255033493041992 - score 0.5506

KEY tp: 2722 - fp: 1965 - fn: 2122 - precision: 0.5808 - recall: 0.5619 - f1-score: 0.5712

KEY tp: 2516 - fp: 1994 - fn: 2050 - precision: 0.5579 - recall: 0.5510 - f1-score: 0.5544
tensor(19.1578, device='cuda:0')

KEY tp: 7592 - fp: 2093 - fn: 2186 - precision: 0.7839 - recall: 0.7764 - f1-score: 0.7801
tensor(10.7877, device='cuda:0')

6:57 / 9:06
EPOCH 57 done: loss 14.7557 - lr 0.0001953
2021-06-28 09:05:36,762 DEV : loss 19.31171226501465 - score 0.5524

KEY tp: 3045 - fp: 2367 - fn: 1799 - precision: 0.5626 - recall: 0.6286 - f1-score: 0.5938

KEY tp: 2740 - fp: 2501 - fn: 1826 - precision: 0.5228 - recall: 0.6001 - f1-score: 0.5588
tensor(19.2284, device='cuda:0')

KEY tp: 6995 - fp: 4207 - fn: 2783 - precision: 0.6244 - recall: 0.7154 - f1-score: 0.6668
tensor(15.8555, device='cuda:0')

6:52 / 8:26
EPOCH 61 done: loss 14.3066 - lr 0.0001953
2021-06-28 08:26:13,793 DEV : loss 19.42903709411621 - score 0.5563

KEY tp: 3041 - fp: 2339 - fn: 1803 - precision: 0.5652 - recall: 0.6278 - f1-score: 0.5949

KEY tp: 2774 - fp: 2429 - fn: 1792 - precision: 0.5332 - recall: 0.6075 - f1-score: 0.5679
tensor(19.1975, device='cuda:0')

KEY tp: 7124 - fp: 3939 - fn: 2654 - precision: 0.6439 - recall: 0.7286 - f1-score: 0.6837
tensor(14.8101, device='cuda:0')

6:58 / 8:18
EPOCH 60 done: loss 14.4531 - lr 0.0001953
2021-06-28 08:18:42,967 DEV : loss 19.19498062133789 - score 0.5506

KEY tp: 2824 - fp: 2080 - fn: 2020 - precision: 0.5759 - recall: 0.5830 - f1-score: 0.5794

KEY tp: 2584 - fp: 2132 - fn: 1982 - precision: 0.5479 - recall: 0.5659 - f1-score: 0.5568
tensor(18.9255, device='cuda:0')

KEY tp: 2584 - fp: 2132 - fn: 1982 - precision: 0.5479 - recall: 0.5659 - f1-score: 0.5568
tensor(18.9255, device='cuda:0')

SEMEVAL 2010

17:38
17:48
51
KEY tp: 297 - fp: 675 - fn: 687 - precision: 0.3056 - recall: 0.3018 - f1-score: 0.3037

17:39
17:49
47
KEY tp: 373 - fp: 1362 - fn: 611 - precision: 0.2150 - recall: 0.3791 - f1-score: 0.2744

17:40
17:52
61

KEY tp: 295 - fp: 472 - fn: 689 - precision: 0.3846 - recall: 0.2998 - f1-score: 0.3370

17:41
17:50
47
KEY tp: 412 - fp: 1234 - fn: 572 - precision: 0.2503 - recall: 0.4187 - f1-score: 0.3133

17:37
17:47
55
KEY tp: 262 - fp: 438 - fn: 722 - precision: 0.3743 - recall: 0.2663 - f1-score: 0.3112

09:59
10:09
52
KEY tp: 322 - fp: 640 - fn: 662 - precision: 0.3347 - recall: 0.3272 - f1-score: 0.3309

09:59
10:09
54
KEY tp: 311 - fp: 734 - fn: 673 - precision: 0.2976 - recall: 0.3161 - f1-score: 0.3066

09:59
10:17
55
KEY tp: 281 - fp: 529 - fn: 703 - precision: 0.3469 - recall: 0.2856 - f1-score: 0.3133

10:00
10:11
61
KEY tp: 300 - fp: 598 - fn: 684 - precision: 0.3341 - recall: 0.3049 - f1-score: 0.3188

09:59
10:10
57
KEY tp: 277 - fp: 527 - fn: 707 - precision: 0.3445 - recall: 0.2815 - f1-score: 0.3098

SEMEVAL2017

1500 S
EPOCH 51 done: loss 32.5930 - lr 0.0001953
2021-06-07 17:37:48,822 DEV : loss 36.1810302734375 - score 0.5638

KEY tp: 901 - fp: 832 - fn: 851 - precision: 0.5199 - recall: 0.5143 - f1-score: 0.5171

EY tp: 542 - fp: 395 - fn: 435 - precision: 0.5784 - recall: 0.5548 - f1-score: 0.5664
tensor(37.2901, device='cuda:0')

EY tp: 2836 - fp: 2307 - fn: 2633 - precision: 0.5514 - recall: 0.5186 - f1-score: 0.5345
tensor(34.7523, device='cuda:0')

2550 S
EPOCH 55 done: loss 30.9846 - lr 0.0001953
2021-06-07 16:54:18,789 DEV : loss 35.963417053222656 - score 0.5694

KEY tp: 1022 - fp: 956 - fn: 730 - precision: 0.5167 - recall: 0.5833 - f1-score: 0.5480

KEY tp: 601 - fp: 471 - fn: 376 - precision: 0.5606 - recall: 0.6151 - f1-score: 0.5866
tensor(37.5122, device='cuda:0')

KEY tp: 3120 - fp: 2654 - fn: 2349 - precision: 0.5404 - recall: 0.5705 - f1-score: 0.5550
tensor(34.6626, device='cuda:0')

1520 s
EPOCH 52 done: loss 32.0662 - lr 0.0001953
2021-06-07 17:38:05,995 DEV : loss 36.758541107177734 - score 0.5596

KEY tp: 915 - fp: 841 - fn: 837 - precision: 0.5211 - recall: 0.5223 - f1-score: 0.5217

EY tp: 564 - fp: 397 - fn: 413 - precision: 0.5869 - recall: 0.5773 - f1-score: 0.5820
tensor(35.6995, device='cuda:0')

EY tp: 3111 - fp: 2106 - fn: 2358 - precision: 0.5963 - recall: 0.5688 - f1-score: 0.5823
tensor(29.1046, device='cuda:0')

1800s
EPOCH 62 done: loss 27.8185 - lr 0.0001953
2021-06-07 17:43:14,950 DEV : loss 36.69933319091797 - score 0.5713

KEY tp: 1008 - fp: 981 - fn: 744 - precision: 0.5068 - recall: 0.5753 - f1-score: 0.5389

EY tp: 598 - fp: 487 - fn: 379 - precision: 0.5512 - recall: 0.6121 - f1-score: 0.5800
tensor(36.4512, device='cuda:0')

EY tp: 3377 - fp: 2503 - fn: 2092 - precision: 0.5743 - recall: 0.6175 - f1-score: 0.5951
tensor(28.6876, device='cuda:0')

2130 s

EPOCH 73 done: loss 25.3282 - lr 0.0001953
2021-06-07 17:49:05,931 DEV : loss 36.47206497192383 - score 0.5727

KEY tp: 939 - fp: 950 - fn: 813 - precision: 0.4971 - recall: 0.5360 - f1-score: 0.5158

EY tp: 593 - fp: 453 - fn: 384 - precision: 0.5669 - recall: 0.6070 - f1-score: 0.5863
tensor(35.4975, device='cuda:0')

EY tp: 3751 - fp: 1852 - fn: 1718 - precision: 0.6695 - recall: 0.6859 - f1-score: 0.6776
tensor(21.5406, device='cuda:0')

2800 s
EPOCH 60 done: loss 32.3727 - lr 0.0001953
2021-06-07 17:49:18,859 DEV : loss 36.179107666015625 - score 0.5726

KEY tp: 884 - fp: 798 - fn: 868 - precision: 0.5256 - recall: 0.5046 - f1-score: 0.5149

EY tp: 549 - fp: 379 - fn: 428 - precision: 0.5916 - recall: 0.5619 - f1-score: 0.5764
tensor(34.7716, device='cuda:0')
EY tp: 3249 - fp: 1909 - fn: 2220 - precision: 0.6299 - recall: 0.5941 - f1-score: 0.6115
tensor(26.0978, device='cuda:0')

2200 s
EPOCH 74 done: loss 25.6365 - lr 0.0001953
2021-06-07 18:32:08,363 DEV : loss 36.50119400024414 - score 0.5758

KEY tp: 884 - fp: 817 - fn: 868 - precision: 0.5197 - recall: 0.5046 - f1-score: 0.5120

EY tp: 557 - fp: 372 - fn: 420 - precision: 0.5996 - recall: 0.5701 - f1-score: 0.5845
tensor(34.6400, device='cuda:0')

EY tp: 3747 - fp: 1415 - fn: 1722 - precision: 0.7259 - recall: 0.6851 - f1-score: 0.7049
tensor(19.3591, device='cuda:0')

1830 s
EPOCH 62 done: loss 29.1650 - lr 0.0001953
2021-06-07 18:25:45,869 DEV : loss 35.70594024658203 - score 0.5859

KEY tp: 943 - fp: 860 - fn: 809 - precision: 0.5230 - recall: 0.5382 - f1-score: 0.5305

EY tp: 591 - fp: 414 - fn: 386 - precision: 0.5881 - recall: 0.6049 - f1-score: 0.5964
tensor(34.5816, device='cuda:0')

EY tp: 3386 - fp: 2060 - fn: 2083 - precision: 0.6217 - recall: 0.6191 - f1-score: 0.6204
tensor(26.2427, device='cuda:0')

1880 s
EPOCH 62 done: loss 29.0049 - lr 0.0001953
2021-06-07 18:22:01,951 DEV : loss 35.87324905395508 - score 0.5653

KEY tp: 922 - fp: 798 - fn: 830 - precision: 0.5360 - recall: 0.5263 - f1-score: 0.5311

EY tp: 556 - fp: 382 - fn: 421 - precision: 0.5928 - recall: 0.5691 - f1-score: 0.5807
tensor(34.8695, device='cuda:0')

EY tp: 3088 - fp: 1995 - fn: 2381 - precision: 0.6075 - recall: 0.5646 - f1-score: 0.5853
tensor(28.7701, device='cuda:0')

1660 s
EPOCH 55 done: loss 30.9344 - lr 0.0001953
2021-06-07 18:22:40,763 DEV : loss 35.96708297729492 - score 0.5723

KEY tp: 1003 - fp: 1050 - fn: 749 - precision: 0.4886 - recall: 0.5725 - f1-score: 0.5272

EY tp: 621 - fp: 509 - fn: 356 - precision: 0.5496 - recall: 0.6356 - f1-score: 0.5895
tensor(37.8917, device='cuda:0')

EY tp: 3223 - fp: 2932 - fn: 2246 - precision: 0.5236 - recall: 0.5893 - f1-score: 0.5545
tensor(33.7368, device='cuda:0')

Experimentos nuevos

BR	1x128		
BRE	2x128		
BR	2x64	4x64	2x256 (LR 0.05 / 0.2)
BR	3x128		
BR	2x128 DO 0.3	3x128 DO 0.3	
B	4x128 DO 0.3	5x64 DO 0.2	
B	3x128 DO 0.3 LR 0.2		
B	Max Batch size 20		
B	Optimizers		

X BERT 2x128 Varios

66
KEY tp: 3024 - fp: 2337 - fn: 1820 - precision: 0.5641 - recall: 0.6243 - f1-score: 0.5927

nKEY tp: 2795 - fp: 2447 - fn: 1771 - precision: 0.5332 - recall: 0.6121 - f1-score: 0.5699
tensor(19.0001, device='cuda:0')
KEY tp: 7217 - fp: 3896 - fn: 2561 - precision: 0.6494 - recall: 0.7381 - f1-score: 0.6909
tensor(14.2928, device='cuda:0')

61
KEY tp: 3118 - fp: 2493 - fn: 1726 - precision: 0.5557 - recall: 0.6437 - f1-score: 0.5965

KEY tp: 2862 - fp: 2524 - fn: 1704 - precision: 0.5314 - recall: 0.6268 - f1-score: 0.5752
tensor(19.8102, device='cuda:0')

nKEY tp: 7887 - fp: 3521 - fn: 1891 - precision: 0.6914 - recall: 0.8066 - f1-score: 0.7445
tensor(12.0780, device='cuda:0')

_ ELMO 2x128

embedding = 'ELMo'
embedding_path = ''
dataset_base_path = ''
dataset = 'Inspec'
output_base_path = '../result/'
iteration = ''
gpu = 1
lr = 0.05
anneal_factor = 0.5
patience = 4
batch_size = 4
num_epochs = 100
threads = 12
param_selection_mode = False
use_tensorboard = False
no_dev = False
use_crf = True
rnn_layers = 2
hidden_size = 128
dropout = 0.0
word_dropout = 0.05
locked_dropout = 0.5
downsample_train = 0.0
not_in_memory = False

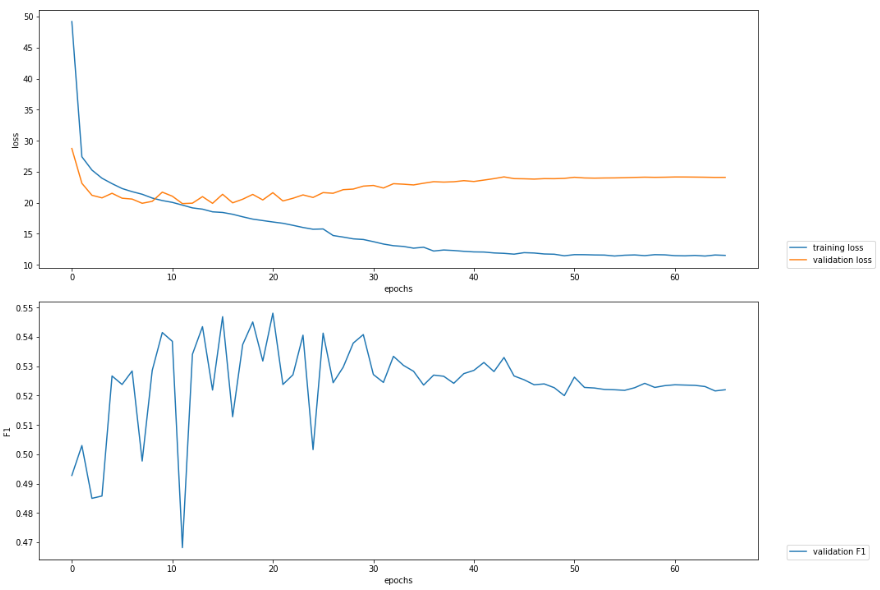
5600 s

EPOCH 66 done: loss 11.5270 - lr 0.0001953
2021-06-30 12:59:10,787 DEV : loss 24.099092483520508 - score 0.522

KEY tp: 3045 - fp: 2732 - fn: 1799 - precision: 0.5271 - recall: 0.6286 - f1-score: 0.5734

KEY tp: 2808 - fp: 2872 - fn: 1758 - precision: 0.4944 - recall: 0.6150 - f1-score: 0.5481
tensor(21.5840, device='cuda:0')

KEY tp: 7675 - fp: 4488 - fn: 2115 - precision: 0.6310 - recall: 0.7840 - f1-score: 0.6992
tensor(14.0255, device='cuda:0')



_ ELMO 1x128

embedding = 'ELMo'

embedding_path = ''
dataset_base_path = ''
dataset = 'Inspec'
output_base_path = '../result/'
iteration = ''
gpu = 1
lr = 0.05
anneal_factor = 0.5
patience = 4
batch_size = 4
num_epochs = 100
threads = 12
param_selection_mode = False
use_tensorboard = False
no_dev = False
use_crf = True
rnn_layers = 1
hidden_size = 128
dropout = 0.0
word_dropout = 0.05
locked_dropout = 0.5
downsample_train = 0.0
not_in_memory = False

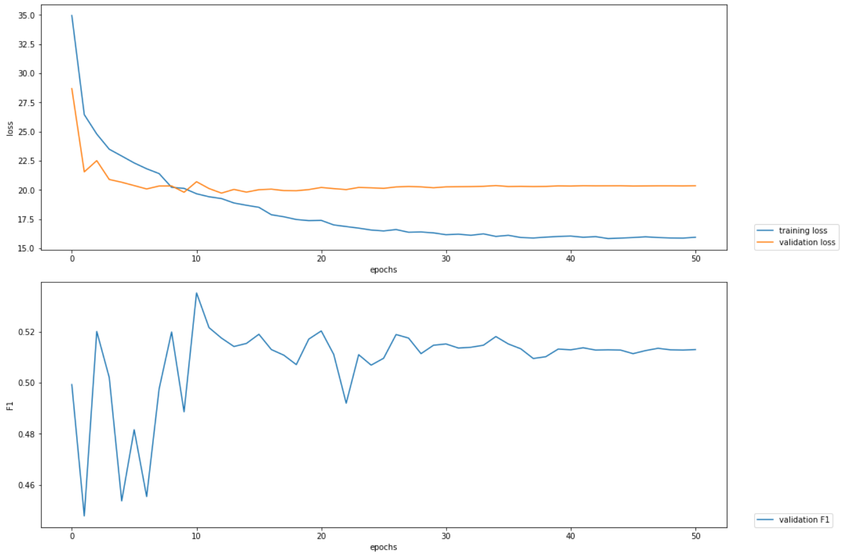
4300 s

EPOCH 51 done: loss 15.9347 - lr 0.0001953
2021-06-30 12:32:31,319 DEV : loss 20.35521125793457 - score 0.513

KEY tp: 2861 - fp: 2539 - fn: 1983 - precision: 0.5298 - recall: 0.5906 - f1-score: 0.5586

KEY tp: 2641 - fp: 2663 - fn: 1925 - precision: 0.4979 - recall: 0.5784 - f1-score: 0.5352
tensor(20.6572, device='cuda:0')

KEY tp: 6705 - fp: 4785 - fn: 3085 - precision: 0.5836 - recall: 0.6849 - f1-score: 0.6302
tensor(17.2631, device='cuda:0')



_ BERT 2x128

embedding = 'Bert'
embedding_path = ''
dataset_base_path = ''
dataset = 'Inspec'
output_base_path = '../result/'
iteration = ''
gpu = 1
lr = 0.05
anneal_factor = 0.5
patience = 4
batch_size = 4
num_epochs = 100
threads = 12
param_selection_mode = False
use_tensorboard = False
no_dev = False
use_crf = True
rnn_layers = 2
hidden_size = 128
dropout = 0.0
word_dropout = 0.05
locked_dropout = 0.5
downsample_train = 0.0
not_in_memory = False

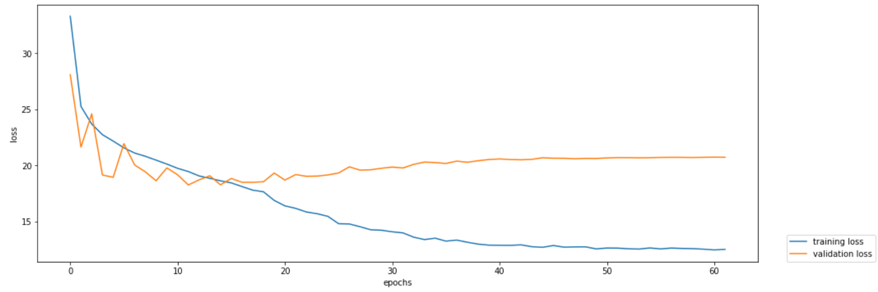
9800 s

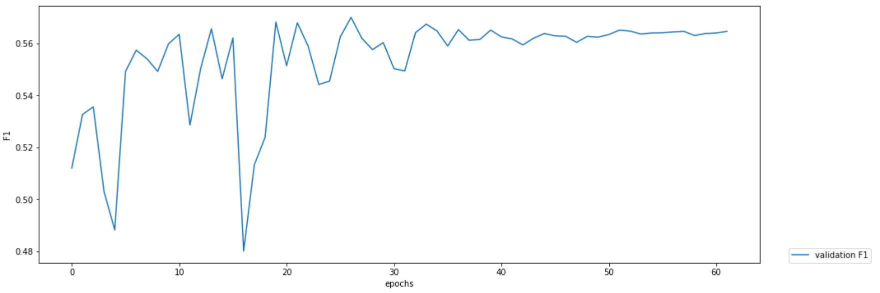
EPOCH 62 done: loss 12.5172 - lr 0.0001953
2021-06-14 11:22:07,973 DEV : loss 20.72060203552246 - score 0.5645

KEY tp: 3044 - fp: 2387 - fn: 1800 - precision: 0.5605 - recall: 0.6284 - f1-score: 0.5925

KEY tp: 2795 - fp: 2447 - fn: 1771 - precision: 0.5332 - recall: 0.6121 - f1-score: 0.5699
tensor(19.8217, device='cuda:0')

KEY tp: 7978 - fp: 3038 - fn: 1800 - precision: 0.7242 - recall: 0.8159 - f1-score: 0.7673
tensor(11.1702, device='cuda:0')





_ BERT 2x128 DO 0.3

embedding = 'Bert'
  embedding_path = ''
  dataset_base_path = ''
  dataset = 'Inspec'
  output_base_path = '../result/'
  iteration = ''
  gpu = 1
  lr = 0.05
  anneal_factor = 0.5
  patience = 4
  batch_size = 4
  num_epochs = 100
  threads = 12
  param_selection_mode = False
  use_tensorboard = False
  no_dev = False
  use_crf = True
  rnn_layers = 2
  hidden_size = 128
  dropout = 0.3
  word_dropout = 0.05
  locked_dropout = 0.5
  downsample_train = 0.0
  not_in_memory = False

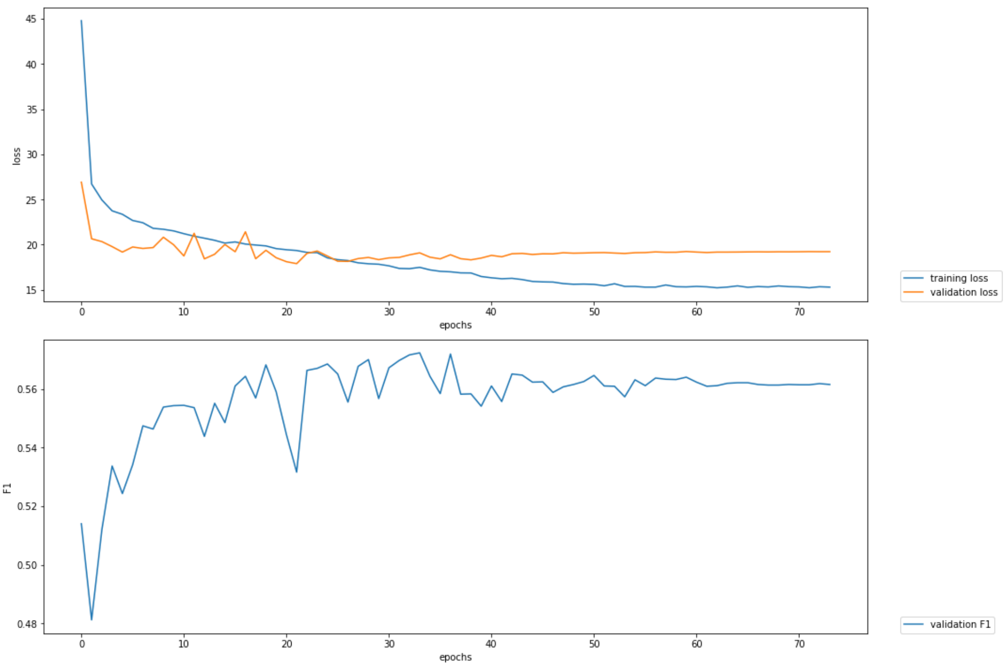
6800 s

EPOCH 74 done: loss 15.3026 - lr 0.0001953
2021-06-14 10:31:38,214 DEV : loss 19.223764419555664 - score 0.5615

KEY tp: 3184 - fp: 2499 - fn: 1660 - precision: 0.5603 - recall: 0.6573 - f1-score: 0.6049

KEY tp: 2872 - fp: 2598 - fn: 1694 - precision: 0.5250 - recall: 0.6290 - f1-score: 0.5723
tensor(19.0466, device='cuda:0')

KEY tp: 7382 - fp: 4352 - fn: 2396 - precision: 0.6291 - recall: 0.7550 - f1-score: 0.6863
tensor(14.6643, device='cuda:0')


loss
epochs
training loss
validation loss
F1
validation F1

_! BERT 3x128 DO 0.3

embedding = 'Bert'
embedding_path = ''
dataset_base_path = ''
dataset = 'Inspec'
output_base_path = '../result/'
iteration = ''
gpu = 1
lr = 0.05
anneal_factor = 0.5
patience = 4
batch_size = 4
num_epochs = 100
threads = 12
param_selection_mode = False
use_tensorboard = False
no_dev = False
use_crf = True
rnn_layers = 3
hidden_size = 128
dropout = 0.3
word_dropout = 0.05
locked_dropout = 0.5
downsample_train = 0.0
not_in_memory = False

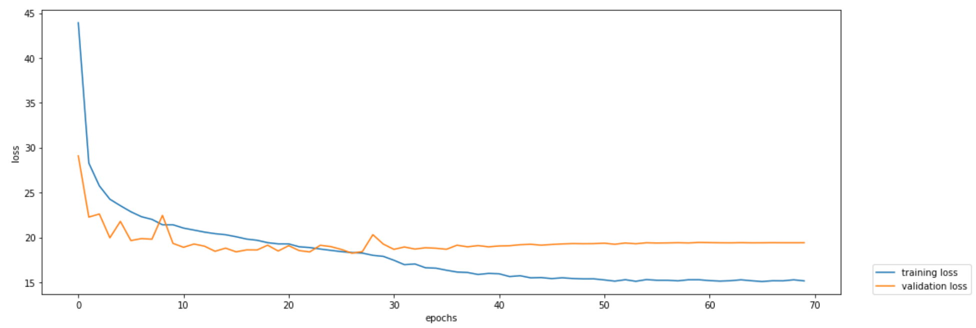
7100 s

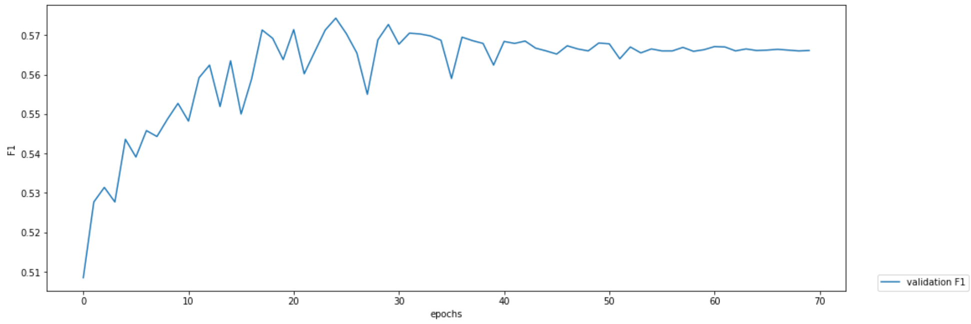
EPOCH 70 done: loss 15.1713 - lr 0.0001953
2021-06-14 10:37:27,874 DEV : loss 19.424774169921875 - score 0.5661

KEY tp: 3190 - fp: 2574 - fn: 1654 - precision: 0.5534 - recall: 0.6585 - f1-score: 0.6014

KEY tp: 2912 - fp: 2663 - fn: 1654 - precision: 0.5223 - recall: 0.6378 - f1-score: 0.5743
tensor(18.9544, device='cuda:0')

KEY tp: 7175 - fp: 4845 - fn: 2603 - precision: 0.5969 - recall: 0.7338 - f1-score: 0.6583
tensor(16.2686, device='cuda:0')





embedding = 'Bert'
embedding_path = ''
dataset_base_path = ''
dataset = 'Inspec'
output_base_path = '../result/'
iteration = ''
gpu = 1
lr = 0.05
anneal_factor = 0.5
patience = 4
batch_size = 16
num_epochs = 100
threads = 12
param_selection_mode = False
use_tensorboard = False
no_dev = False
use_crf = True
rnn_layers = 3
hidden_size = 128
dropout = 0.3
word_dropout = 0.05
locked_dropout = 0.5
downsample_train = 0.0
not_in_memory = False

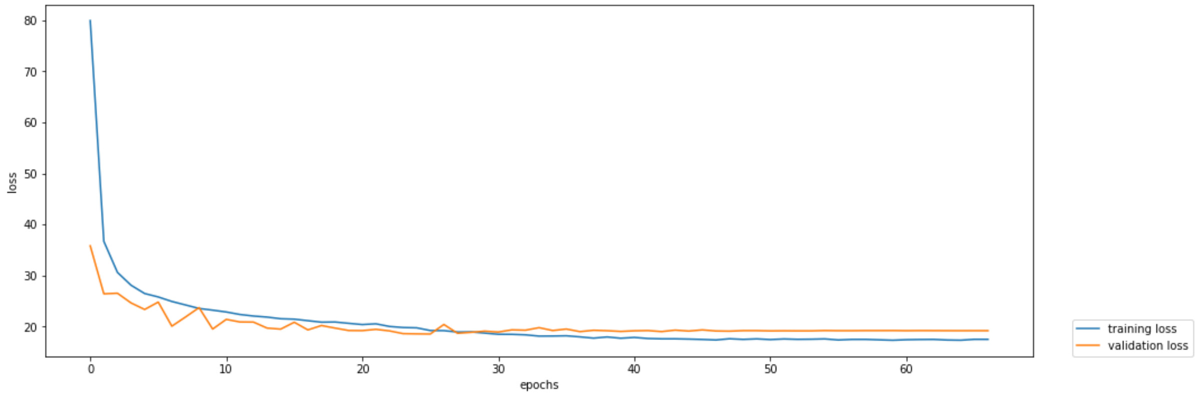
4500 s

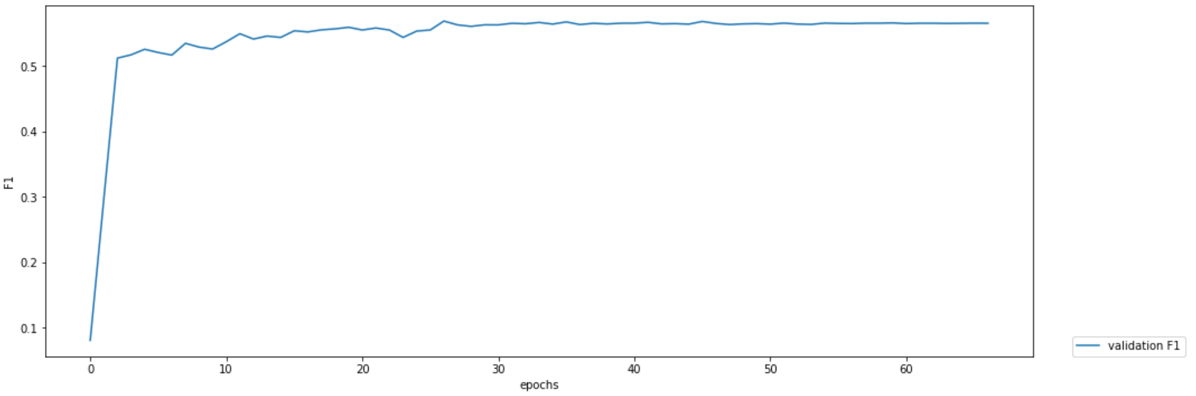
EPOCH 67 done: loss 17.4535 - lr 0.0001953
2021-06-17 17:16:15,374 DEV : loss 19.14909553527832 - score 0.5657

KEY tp: 3414 - fp: 3137 - fn: 1430 - precision: 0.5211 - recall: 0.7048 - f1-score: 0.5992

KEY tp: 3095 - fp: 3218 - fn: 1471 - precision: 0.4903 - recall: 0.6778 - f1-score: 0.5690
tensor(20.3325, device='cuda:0')

KEY tp: 7376 - fp: 6053 - fn: 2402 - precision: 0.5493 - recall: 0.7543 - f1-score: 0.6357
tensor(18.1900, device='cuda:0')





embedding = 'Bert'
embedding_path = ''
dataset_base_path = ''
dataset = 'Inspec'
output_base_path = '../result/'
iteration = ''
gpu = 1
lr = 0.05
anneal_factor = 0.5
patience = 4
batch_size = 4
num_epochs = 100
threads = 12
param_selection_mode = False
use_tensorboard = False
no_dev = False
use_crf = True
rnn_layers = 3
hidden_size = 128
dropout = 0.3
word_dropout = 0.05
locked_dropout = 0.5
downsample_train = 0.0
not_in_memory = False

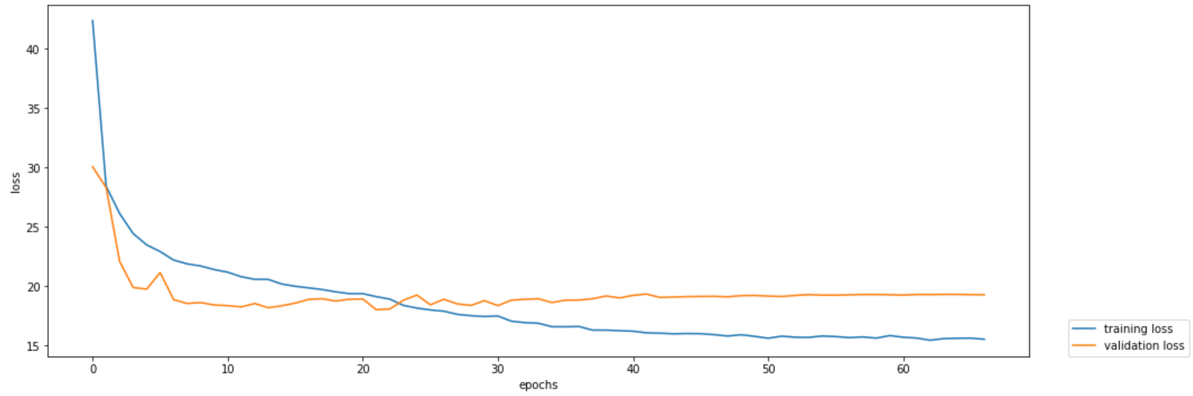
13000 s

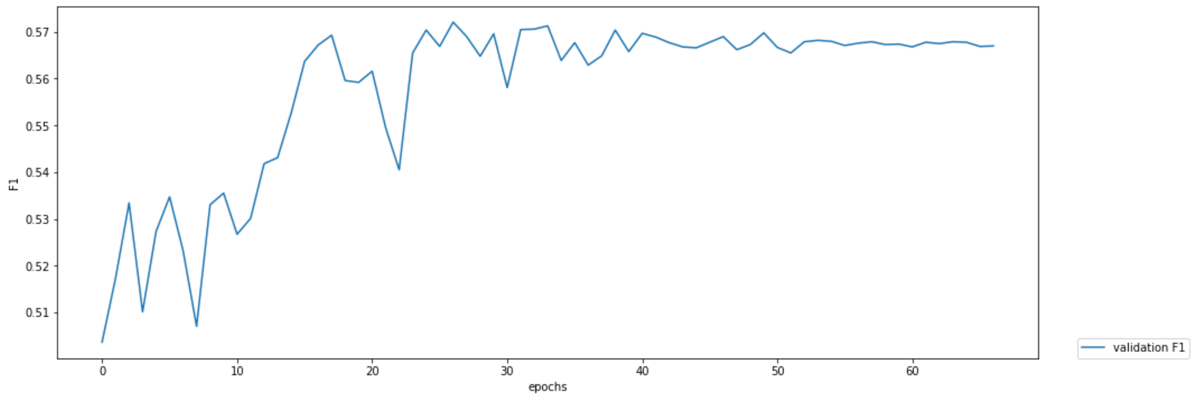
EPOCH 67 done: loss 15.5349 - lr 0.0001953
2021-06-17 19:53:11,058 DEV : loss 19.28322410583496 - score 0.567

KEY tp: 3130 - fp: 2516 - fn: 1714 - precision: 0.5544 - recall: 0.6462 - f1-score: 0.5968

KEY tp: 2861 - fp: 2574 - fn: 1705 - precision: 0.5264 - recall: 0.6266 - f1-score: 0.5721
tensor(18.8586, device='cuda:0')

KEY tp: 7336 - fp: 4347 - fn: 2442 - precision: 0.6279 - recall: 0.7503 - f1-score: 0.6837
tensor(15.3139, device='cuda:0')





_ ROBERTA 2x128

embedding = 'RoBERTa'
embedding_path = ''
dataset_base_path = ''
dataset = 'Inspec'
output_base_path = '../result/'
iteration = ''
gpu = 1
lr = 0.05
anneal_factor = 0.5
patience = 4
batch_size = 4
num_epochs = 100
threads = 12
param_selection_mode = False
use_tensorboard = False
no_dev = False
use_crf = True
rnn_layers = 2
hidden_size = 128
dropout = 0.0
word_dropout = 0.05
locked_dropout = 0.5
downsample_train = 0.0
not_in_memory = False

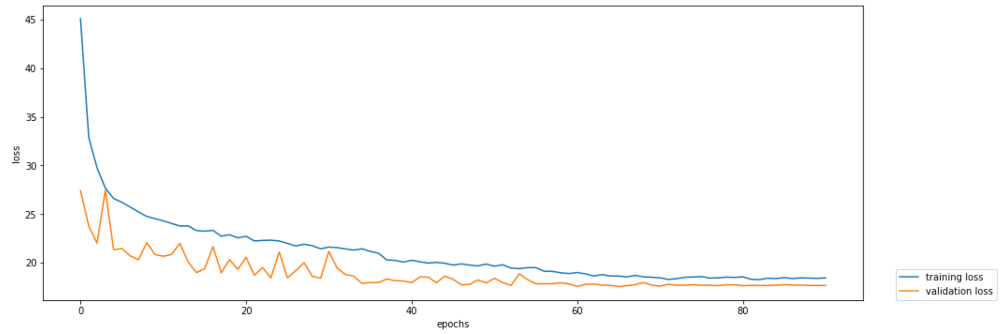
13800 s

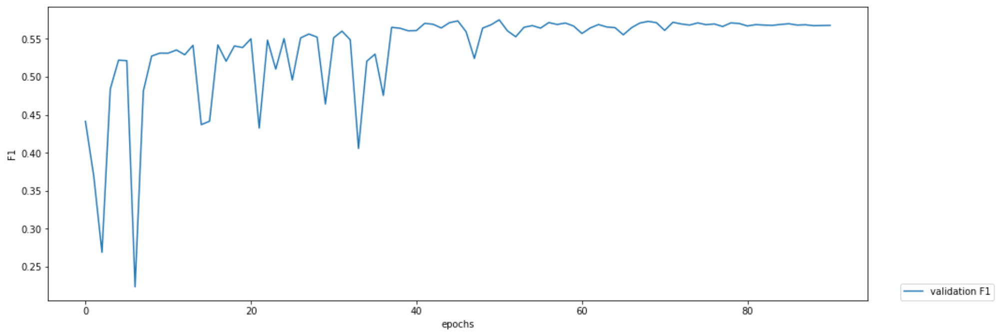
EPOCH 91 done: loss 18.4340 - lr 0.0001953
2021-06-14 12:31:18,829 DEV : loss 17.655641555786133 - score 0.5677

KEY	tp: 3115 - fp: 2587 - fn: 1729 - precision: 0.5463 - recall: 0.6431 - f1-score: 0.5907

KEY	tp: 2914 - fp: 2656 - fn: 1652 - precision: 0.5232 - recall: 0.6382 - f1-score: 0.5750
tensor(18.3328, device='cuda:0')

KEY	tp: 6764 - fp: 5142 - fn: 3017 - precision: 0.5681 - recall: 0.6915 - f1-score: 0.6238
tensor(18.1138, device='cuda:0')





embedding = 'RoBERTa'
embedding_path = ''
dataset_base_path = ''
dataset = 'Inspec'
output_base_path = '../result/'
iteration = ''
gpu = 1
lr = 0.05
anneal_factor = 0.5
patience = 4
batch_size = 16
num_epochs = 100
threads = 12
param_selection_mode = False
use_tensorboard = False
no_dev = False
use_crf = True
rnn_layers = 2
hidden_size = 128
dropout = 0.0
word_dropout = 0.05
locked_dropout = 0.5
downsample_train = 0.0
not_in_memory = False

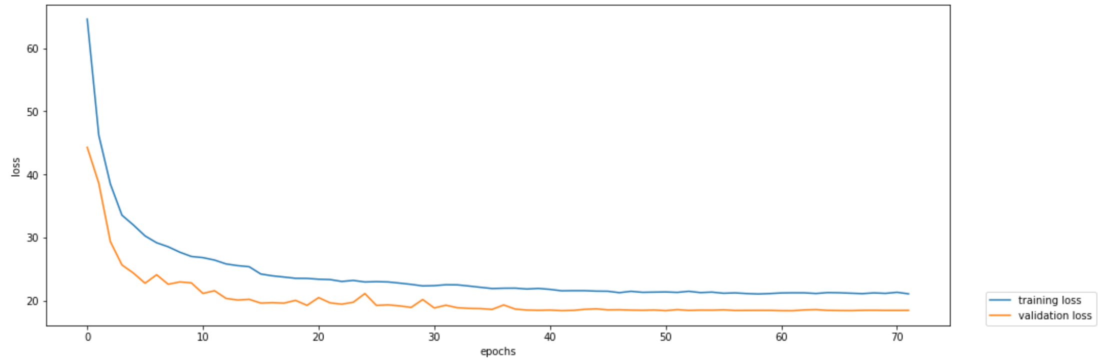
3600 s

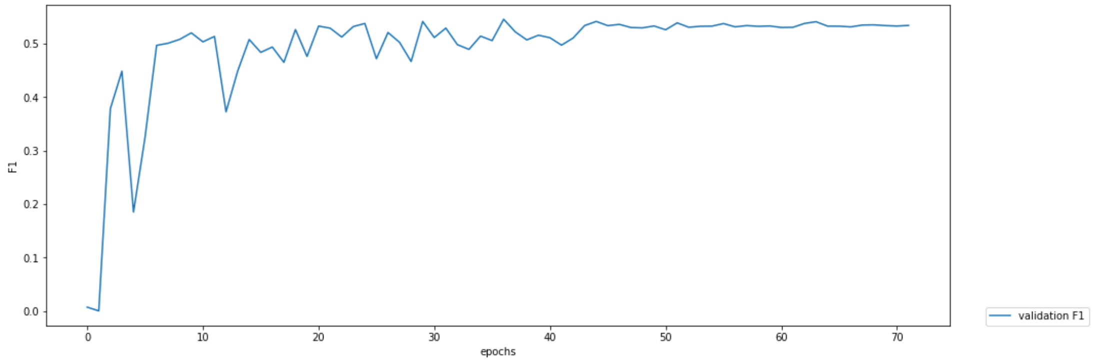
EPOCH 72 done: loss 21.0909 - lr 0.0001953
2021-06-17 14:44:27,474 DEV : loss 18.481929779052734 - score 0.5341

KEY tp: 2957 - fp: 2431 - fn: 1887 - precision: 0.5488 - recall: 0.6104 - f1-score: 0.5780

KEY tp: 2706 - fp: 2646 - fn: 1860 - precision: 0.5056 - recall: 0.5926 - f1-score: 0.5457
tensor(19.3162, device='cuda:0')

KEY tp: 5954 - fp: 5342 - fn: 3827 - precision: 0.5271 - recall: 0.6087 - f1-score: 0.5650
tensor(20.2687, device='cuda:0')





embedding = 'RoBERTa'
embedding_path = ''
dataset_base_path = ''
dataset = 'Inspec'
output_base_path = '../result/'
iteration = ''
gpu = 1
lr = 0.05
anneal_factor = 0.5
patience = 4
batch_size = 4
num_epochs = 100
threads = 12
param_selection_mode = False
use_tensorboard = False
no_dev = False
use_crf = True
rnn_layers = 2
hidden_size = 128
dropout = 0.0
word_dropout = 0.05
locked_dropout = 0.5
downsample_train = 0.0
not_in_memory = False

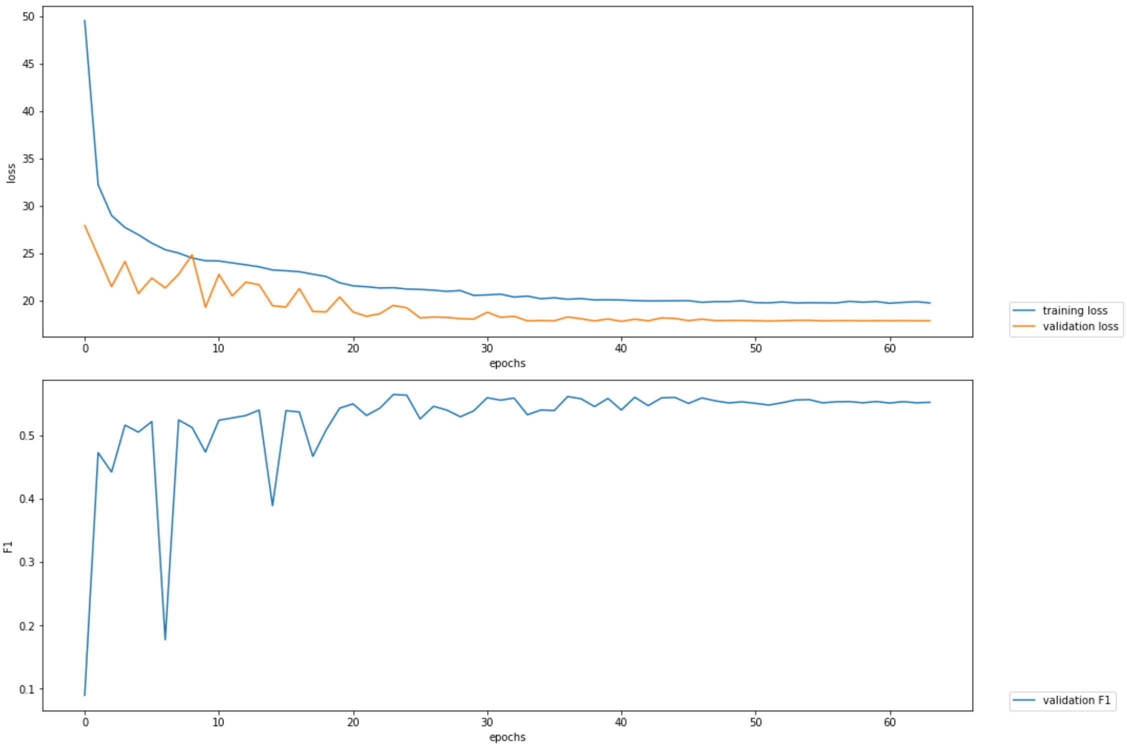
5200 s

EPOCH 64 done: loss 19.7294 - lr 0.0001953
2021-06-17 17:00:33,132 DEV : loss 17.85202980041504 - score 0.5526

KEY tp: 3249 - fp: 2886 - fn: 1595 - precision: 0.5296 - recall: 0.6707 - f1-score: 0.5919

KEY tp: 3007 - fp: 3075 - fn: 1559 - precision: 0.4944 - recall: 0.6586 - f1-score: 0.564
tensor(19.4188, device='cuda:0')

KEY tp: 6703 - fp: 6230 - fn: 3078 - precision: 0.5183 - recall: 0.6853 - f1-score: 0.5902
tensor(20.1764, device='cuda:0')



embedding = 'RoBERTa'
embedding_path = ''
dataset_base_path = ''
dataset = 'Inspec'
output_base_path = '../result/'
iteration = ''
gpu = 1
lr = 0.05
anneal_factor = 0.5
patience = 4
batch_size = 4
num_epochs = 100
threads = 12
param_selection_mode = False
use_tensorboard = False
no_dev = False
use_crf = True
rnn_layers = 2
hidden_size = 128
dropout = 0.0
word_dropout = 0.05
locked_dropout = 0.5
downsample_train = 0.0
not_in_memory = False

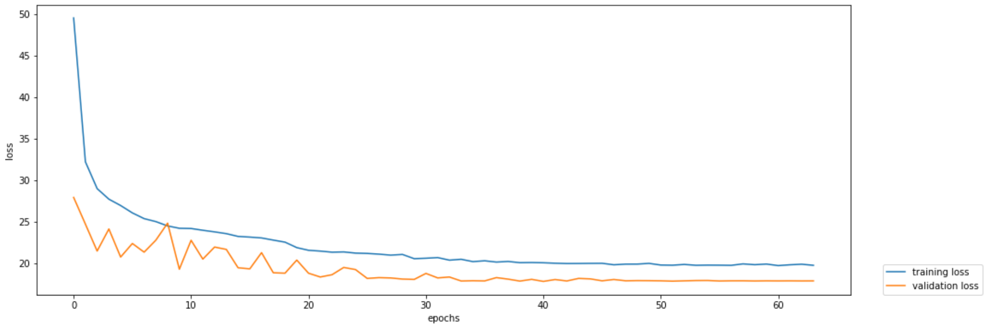
5200 s

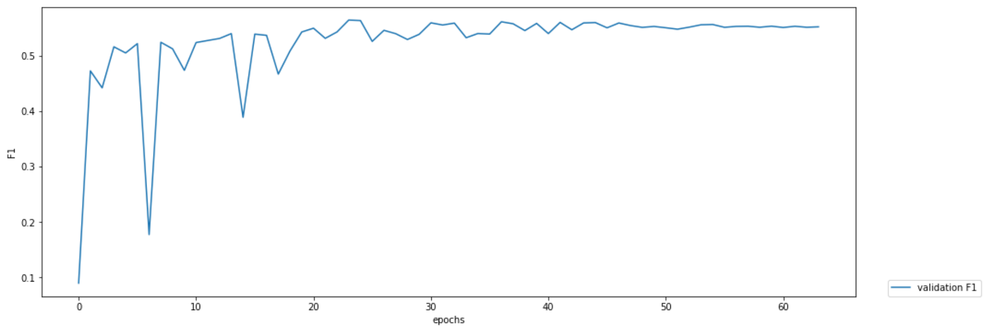
EPOCH 64 done: loss 19.7294 - lr 0.0001953
2021-06-17 17:00:33,132 DEV : loss 17.85202980041504 - score 0.5526

KEY tp: 3249 - fp: 2886 - fn: 1595 - precision: 0.5296 - recall: 0.6707 - f1-score: 0.5919

KEY tp: 3007 - fp: 3075 - fn: 1559 - precision: 0.4944 - recall: 0.6586 - f1-score: 0.5648
tensor(19.4188, device='cuda:0')

KEY tp: 6703 - fp: 6230 - fn: 3078 - precision: 0.5183 - recall: 0.6853 - f1-score: 0.5902
tensor(20.1764, device='cuda:0')





_ ROBERTA 3x128

embedding = 'RoBERTa'
embedding_path = ''
dataset_base_path = ''
dataset = 'Inspec'
output_base_path = '../result/'
iteration = ''
gpu = 1
lr = 0.05
anneal_factor = 0.5
patience = 4
batch_size = 4
num_epochs = 100
threads = 12
param_selection_mode = False
use_tensorboard = False
no_dev = False
use_crf = True
rnn_layers = 3
hidden_size = 128
dropout = 0.0
word_dropout = 0.05
locked_dropout = 0.5
downsample_train = 0.0
not_in_memory = False

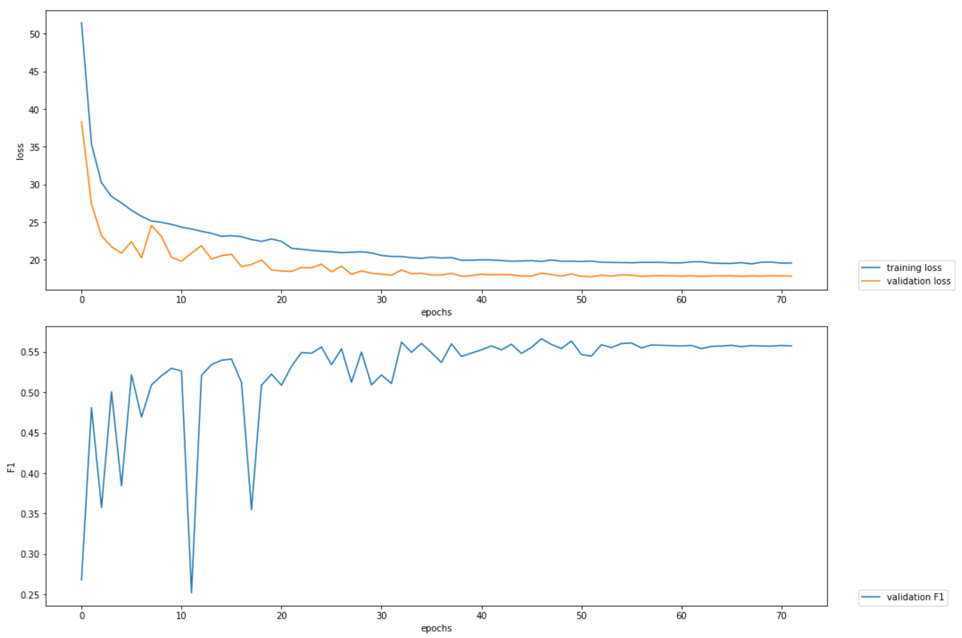
6200 s

EPOCH 71 done: loss 19.5664 - lr 0.0001953
2021-06-14 10:25:42,680 DEV : loss 17.879884719848633 - score 0.5581

KEY tp: 3040 - fp: 2360 - fn: 1804 - precision: 0.5630 - recall: 0.6276 - f1-score: 0.5935

KEY tp: 2800 - fp: 2521 - fn: 1766 - precision: 0.5262 - recall: 0.6132 - f1-score: 0.5664
tensor(18.1923, device='cuda:0')

KEY tp: 6468 - fp: 4847 - fn: 3313 - precision: 0.5716 - recall: 0.6613 - f1-score: 0.6132
tensor(18.3694, device='cuda:0')



_ ROBERTA 1x128

embedding = 'RoBERTa'
  embedding_path = ''
  dataset_base_path = ''
  dataset = 'Inspec'
  output_base_path = '../result/'
  iteration = ''
  gpu = 1
  lr = 0.05
  anneal_factor = 0.5
  patience = 4
  batch_size = 4
  num_epochs = 100
  threads = 12
  param_selection_mode = False
  use_tensorboard = False
  no_dev = False
  use_crf = True
  rnn_layers = 1
  hidden_size = 128
  dropout = 0.0
  word_dropout = 0.05
  locked_dropout = 0.5
  downsample_train = 0.0
  not_in_memory = False

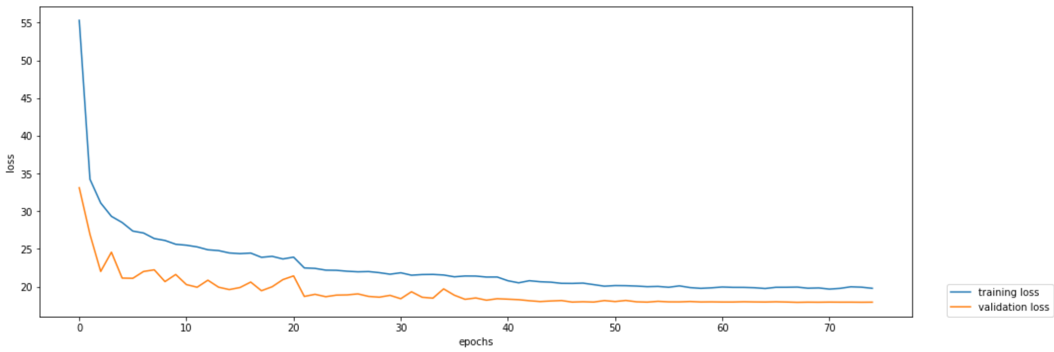
5500 s

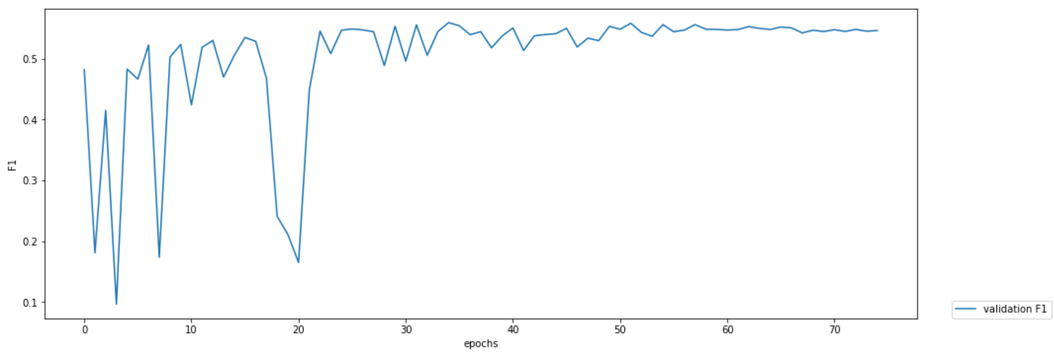
EPOCH 75 done: loss 19.7785 - lr 0.0001953
2021-06-14 14:17:34,410 DEV : loss 17.934602737426758 - score 0.5461

KEY   tp: 3203 - fp: 2767 - fn: 1641 - precision: 0.5365 - recall: 0.6612 - f1-score: 0.5924

KEY   tp: 2925 - fp: 2970 - fn: 1641 - precision: 0.4962 - recall: 0.6406 - f1-score: 0.5592
tensor(19.6413, device='cuda:0')

KEY   tp: 6636 - fp: 6091 - fn: 3145 - precision: 0.5214 - recall: 0.6785 - f1-score: 0.5897
tensor(20.3495, device='cuda:0')





_ BERT 1x128

embedding = 'Bert'
embedding_path = ''
dataset_base_path = ''
dataset = 'Inspec'
output_base_path = '../result/'
iteration = ''
gpu = 1
lr = 0.05
anneal_factor = 0.5
patience = 4
batch_size = 4
num_epochs = 100
threads = 12
param_selection_mode = False
use_tensorboard = False
no_dev = False
use_crf = True
rnn_layers = 1
hidden_size = 128
dropout = 0.0
word_dropout = 0.05
locked_dropout = 0.5
downsample_train = 0.0
not_in_memory = False

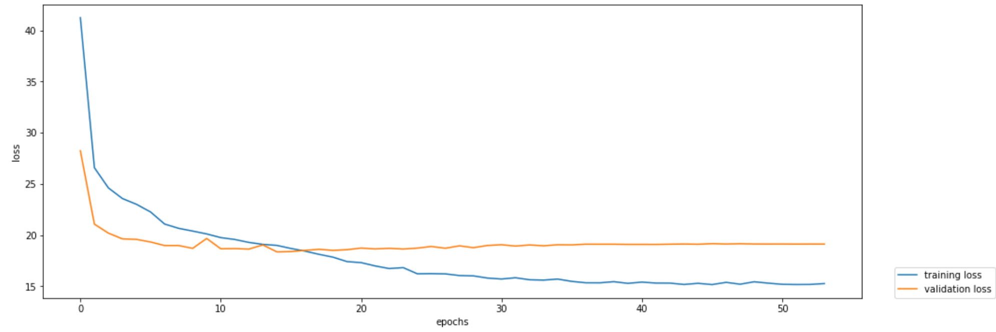
xxxx s

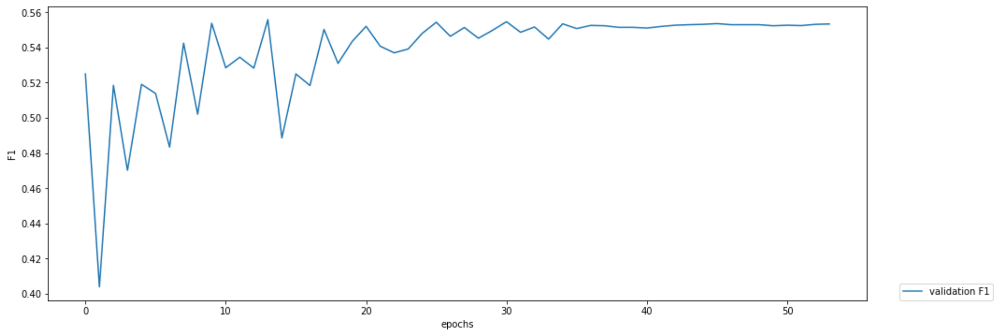
EPOCH 54 done: loss 15.2627 - lr 0.0001953
2021-06-15 09:12:12,590 DEV : loss 19.11808967590332 - score 0.5534

KEY tp: 2993 - fp: 2277 - fn: 1851 - precision: 0.5679 - recall: 0.6179 - f1-score: 0.5919

KEY tp: 2703 - fp: 2456 - fn: 1863 - precision: 0.5239 - recall: 0.5920 - f1-score: 0.5559
tensor(18.9950, device='cuda:0')

KEY tp: 6711 - fp: 4251 - fn: 3067 - precision: 0.6122 - recall: 0.6863 - f1-score: 0.6472
tensor(16.5383, device='cuda:0')





_ ROBERTA 2x128 0.3 DO

embedding = 'RoBERTa'
embedding_path = ''
dataset_base_path = ''
dataset = 'Inspec'
output_base_path = '../result/'
iteration = ''
gpu = 1
lr = 0.05
anneal_factor = 0.5
patience = 4
batch_size = 4
num_epochs = 100
threads = 12
param_selection_mode = False
use_tensorboard = False
no_dev = False
use_crf = True
rnn_layers = 2
hidden_size = 128
dropout = 0.3
word_dropout = 0.05
locked_dropout = 0.5
downsample_train = 0.0
not_in_memory = False

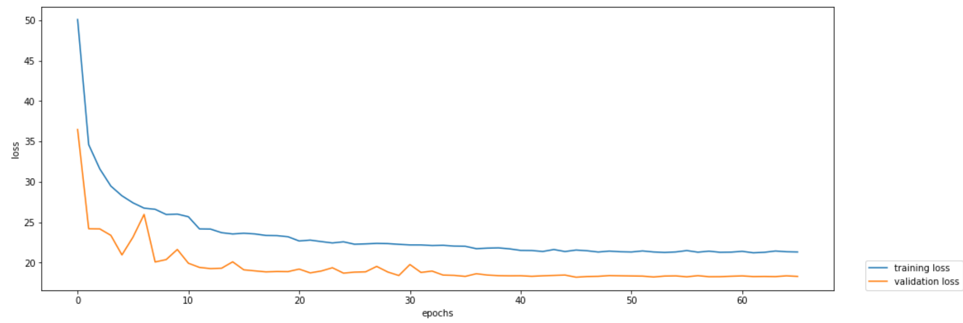
5800 s

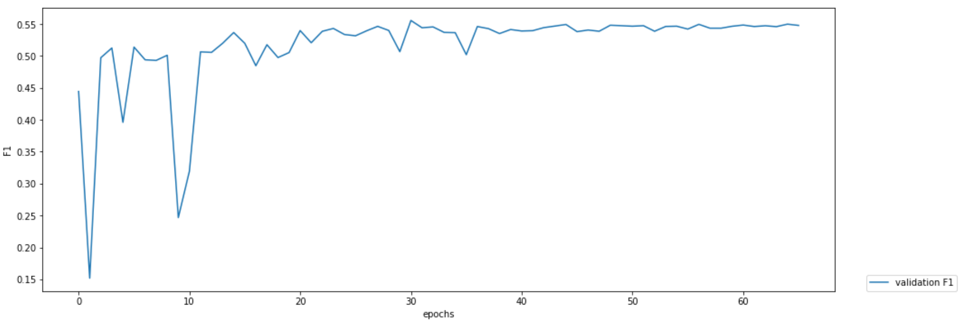
EPOCH 66 done: loss 21.3211 - lr 0.0001953
2021-06-15 10:02:39,533 DEV : loss 18.299531936645508 - score 0.5479

KEY tp: 3206 - fp: 2855 - fn: 1638 - precision: 0.5290 - recall: 0.6618 - f1-score: 0.5880

KEY tp: 2933 - fp: 3058 - fn: 1633 - precision: 0.4896 - recall: 0.6424 - f1-score: 0.5557
tensor(19.7252, device='cuda:0')

KEY tp: 6571 - fp: 6178 - fn: 3210 - precision: 0.5154 - recall: 0.6718 - f1-score: 0.5833
tensor(20.7678, device='cuda:0')





_ ROBERTA 3x128 0.3 DO

embedding = 'RoBERTa'
  embedding_path = ''
  dataset_base_path = ''
  dataset = 'Inspec'
  output_base_path = '../result/'
  iteration = ''
  gpu = 1
  lr = 0.05
  anneal_factor = 0.5
  patience = 4
  batch_size = 4
  num_epochs = 100
  threads = 12
  param_selection_mode = False
  use_tensorboard = False
  no_dev = False
  use_crf = True
  rnn_layers = 3
  hidden_size = 128
  dropout = 0.3
  word_dropout = 0.05
  locked_dropout = 0.5
  downsample_train = 0.0
  not_in_memory = False

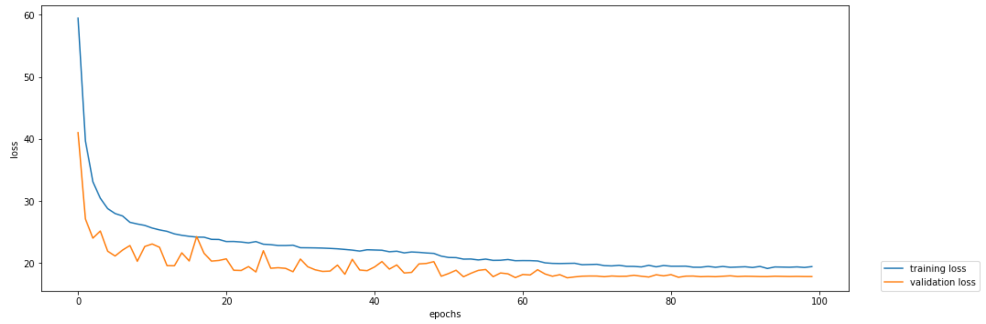
8400 s

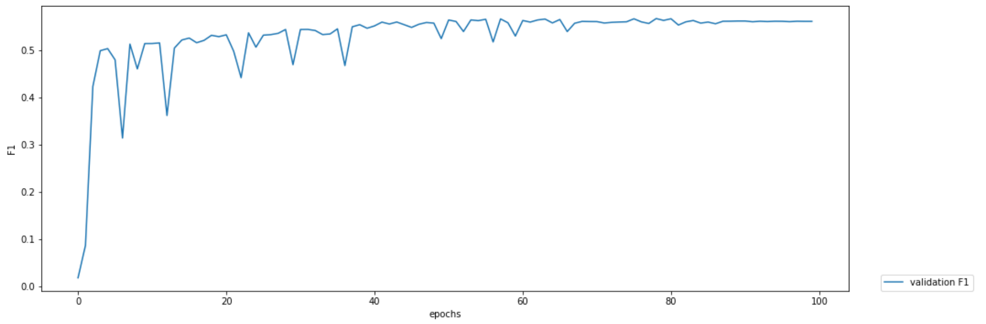
EPOCH 100 done: loss 19.4401 - lr 0.0001953
2021-06-15 11:44:40,805 DEV : loss 17.839059829711914 - score 0.5613

KEY tp: 3040 - fp: 2332 - fn: 1804 - precision: 0.5659 - recall: 0.6276 - f1-score: 0.5951

KEY tp: 2808 - fp: 2528 - fn: 1758 - precision: 0.5262 - recall: 0.6150 - f1-score: 0.5672
tensor(18.0726, device='cuda:0')

KEY tp: 6656 - fp: 4730 - fn: 3125 - precision: 0.5846 - recall: 0.6805 - f1-score: 0.6289
tensor(17.5651, device='cuda:0')





X BERT 5x64 DO 0.2

embedding = 'Bert'
embedding_path = ''
dataset_base_path = ''
dataset = 'Inspec'
output_base_path = '../result/'
iteration = ''
gpu = 1
lr = 0.2
anneal_factor = 0.5
patience = 4
batch_size = 4
num_epochs = 100
threads = 12
param_selection_mode = False
use_tensorboard = False
no_dev = False
use_crf = True
rnn_layers = 5
hidden_size = 64
dropout = 0.2
word_dropout = 0.05
locked_dropout = 0.5
downsample_train = 0.0
not_in_memory = False

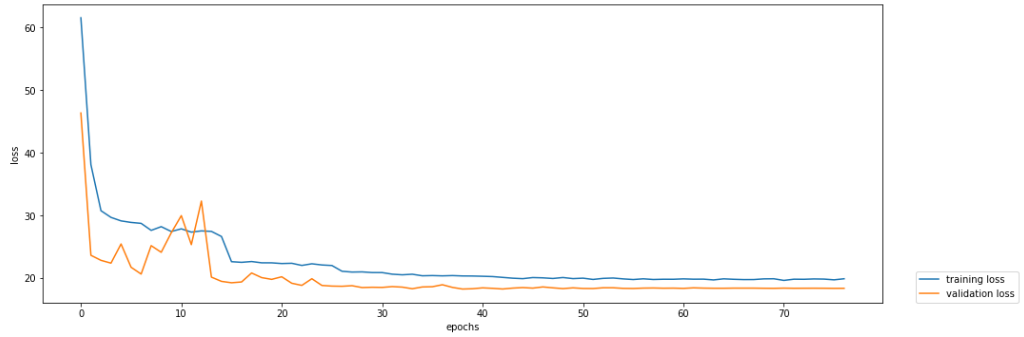
7500 s

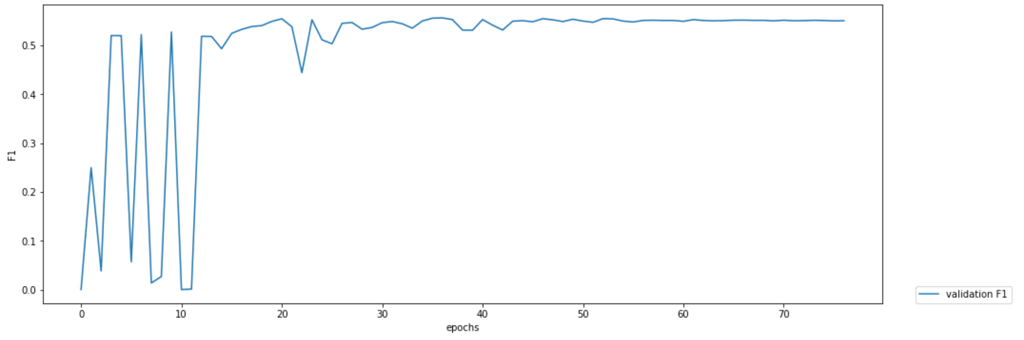
EPOCH 77 done: loss 19.8987 - lr 0.0001953
2021-06-16 13:38:59,334 DEV : loss 18.3765926361084 - score 0.5504

KEY tp: 3121 - fp: 2603 - fn: 1723 - precision: 0.5452 - recall: 0.6443 - f1-score: 0.5907

KEY tp: 2827 - fp: 2774 - fn: 1739 - precision: 0.5047 - recall: 0.6191 - f1-score: 0.5561
tensor(18.9073, device='cuda:0')

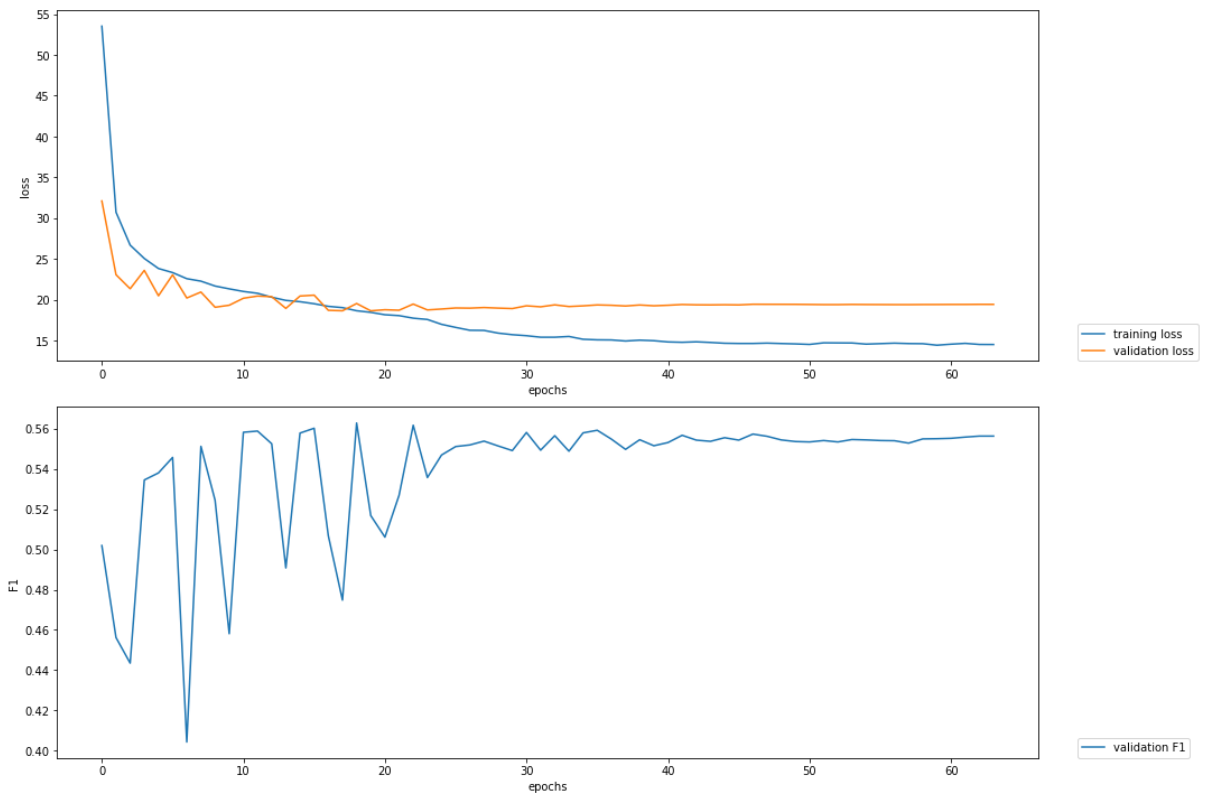
KEY tp: 6477 - fp: 5445 - fn: 3301 - precision: 0.5433 - recall: 0.6624 - f1-score: 0.5970
tensor(19.1130, device='cuda:0')





BATCH SIZE Def

20
2000 s
KEY tp: 3125 - fp: 2549 - fn: 1719 - precision: 0.5508 - recall: 0.6451 - f1-score: 0.5942



X BERT 5x128 DO 0.2 LR 0.2

embedding = 'Bert'
 embedding_path = ''
 dataset_base_path = ''
 dataset = 'Inspec'
 output_base_path = '../result/'
 iteration = ''
 gpu = 1
 lr = 0.2
 anneal_factor = 0.5
 patience = 4
 batch_size = 4
 num_epochs = 100
 threads = 12
 param_selection_mode = False
 use_tensorboard = False
 no_dev = False
 use_crf = True
 rnn_layers = 5
 hidden_size = 128
 dropout = 0.2
 word_dropout = 0.05
 locked_dropout = 0.5
 downsample_train = 0.0
 not_in_memory = False

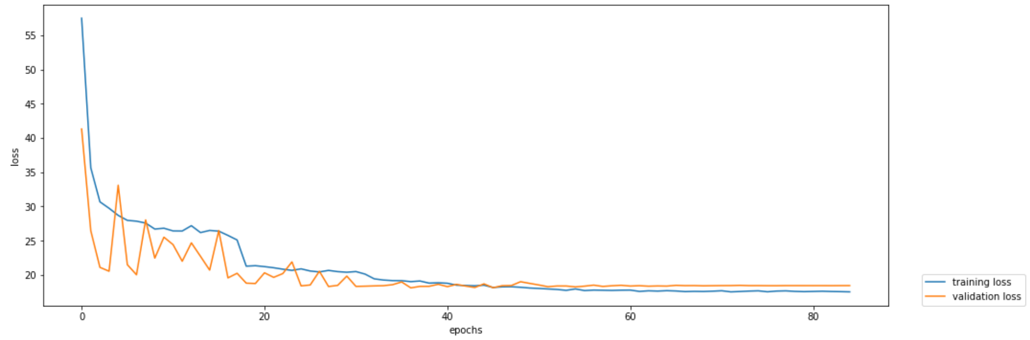
8300 s

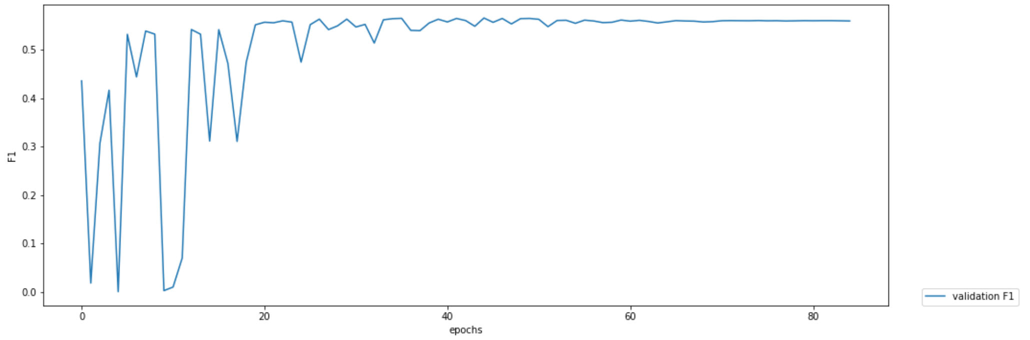
EPOCH 85 done: loss 17.5268 - lr 0.0001953
2021-06-16 17:17:34,354 DEV : loss 18.441688537597656 - score 0.5593

KEY tp: 3141 - fp: 2509 - fn: 1703 - precision: 0.5559 - recall: 0.6484 - f1-score: 0.5986

KEY tp: 2854 - fp: 2676 - fn: 1712 - precision: 0.5161 - recall: 0.6251 - f1-score: 0.5654
tensor(18.6466, device='cuda:0')

KEY tp: 7048 - fp: 4951 - fn: 2730 - precision: 0.5874 - recall: 0.7208 - f1-score: 0.6473
tensor(16.8055, device='cuda:0')





X ROBERTA 5x128 DO 0.2 LR 0.2

embedding = 'RoBERTa'
embedding_path = ''
dataset_base_path = ''
dataset = 'Inspec'
output_base_path = '../result/'
iteration = ''
gpu = 1
lr = 0.2
anneal_factor = 0.5
patience = 4
batch_size = 4
num_epochs = 100
threads = 12
param_selection_mode = False
use_tensorboard = False
no_dev = False
use_crf = True
rnn_layers = 5
hidden_size = 128
dropout = 0.2
word_dropout = 0.05
locked_dropout = 0.5
downsample_train = 0.0
not_in_memory = False

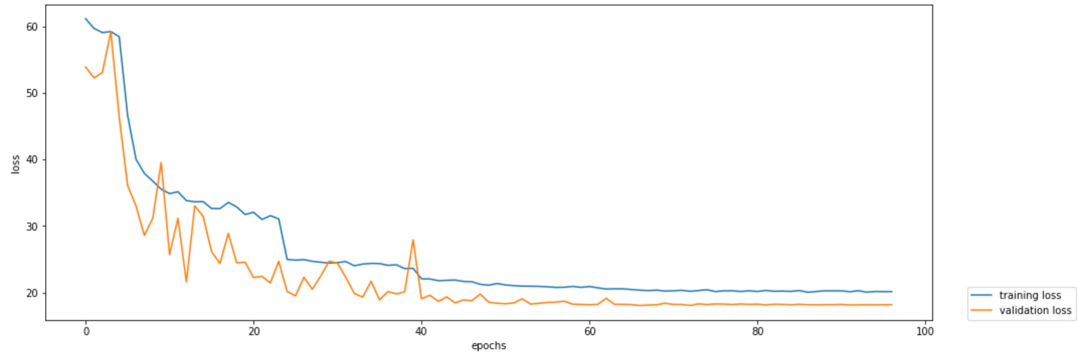
9500 s

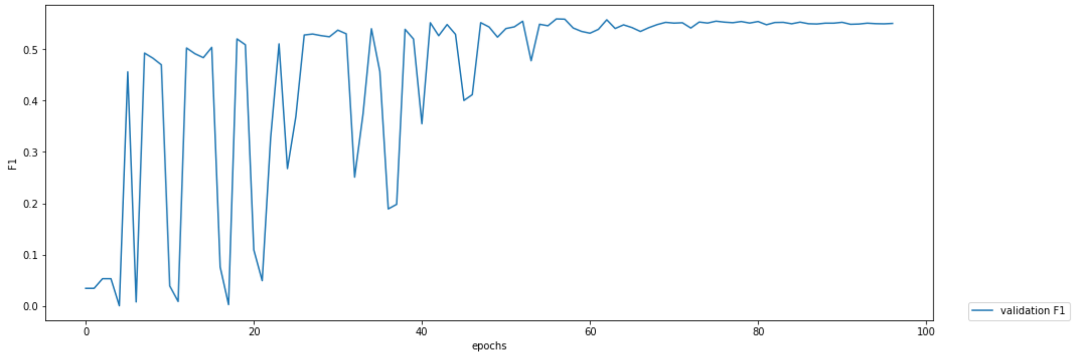
EPOCH 97 done: loss 20.1384 - lr 0.0001953
2021-06-16 20:56:56,585 DEV : loss 18.168703079223633 - score 0.5508

KEY tp: 2950 - fp: 2391 - fn: 1894 - precision: 0.5523 - recall: 0.6090 - f1-score: 0.5793

KEY tp: 2755 - fp: 2526 - fn: 1811 - precision: 0.5217 - recall: 0.6034 - f1-score: 0.5596
tensor(18.5011, device='cuda:0')

KEY tp: 6226 - fp: 4984 - fn: 3555 - precision: 0.5554 - recall: 0.6365 - f1-score: 0.5932
tensor(18.8450, device='cuda:0')





_ ROBERTA 2x256

embedding = 'RoBERTa'
embedding_path = ''
dataset_base_path = ''
dataset = 'Inspec'
output_base_path = '../result/'
iteration = ''
gpu = 1
lr = 0.05
anneal_factor = 0.5
patience = 4
batch_size = 16
num_epochs = 100
threads = 12
param_selection_mode = False
use_tensorboard = False
no_dev = False
use_crf = True
rnn_layers = 2
hidden_size = 256
dropout = 0.0
word_dropout = 0.05
locked_dropout = 0.5
downsample_train = 0.0
not_in_memory = False

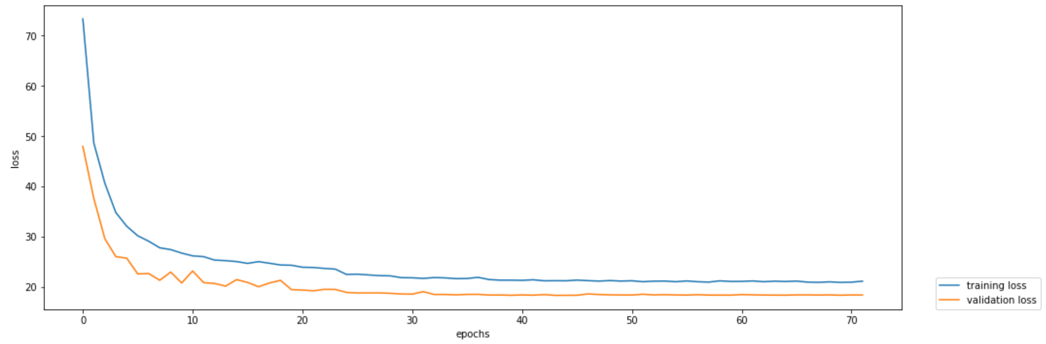
8500 s

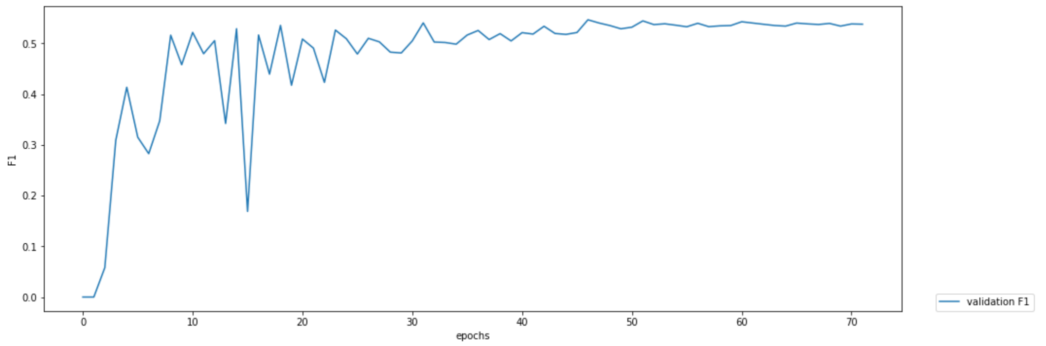
EPOCH 72 done: loss 21.1546 - lr 0.0001953
2021-06-17 12:18:55,281 DEV : loss 18.40152359008789 - score 0.538

KEY tp: 2729 - fp: 2024 - fn: 2115 - precision: 0.5742 - recall: 0.5634 - f1-score: 0.5687

KEY tp: 2524 - fp: 2143 - fn: 2042 - precision: 0.5408 - recall: 0.5528 - f1-score: 0.5467
tensor(18.6009, device='cuda:0')

KEY tp: 5626 - fp: 4253 - fn: 4155 - precision: 0.5695 - recall: 0.5752 - f1-score: 0.5723
tensor(19.4132, device='cuda:0')





_ ROBERTA 2x256 LR 0.2

embedding = 'RoBERTa'
embedding_path = ''
dataset_base_path = ''
dataset = 'Inspec'
output_base_path = '../result/'
iteration = ''
gpu = 1
lr = 0.2
anneal_factor = 0.5
patience = 4
batch_size = 16
num_epochs = 100
threads = 12
param_selection_mode = False
use_tensorboard = False
no_dev = False
use_crf = True
rnn_layers = 2
hidden_size = 256
dropout = 0.0
word_dropout = 0.05
locked_dropout = 0.5
downsample_train = 0.0
not_in_memory = False

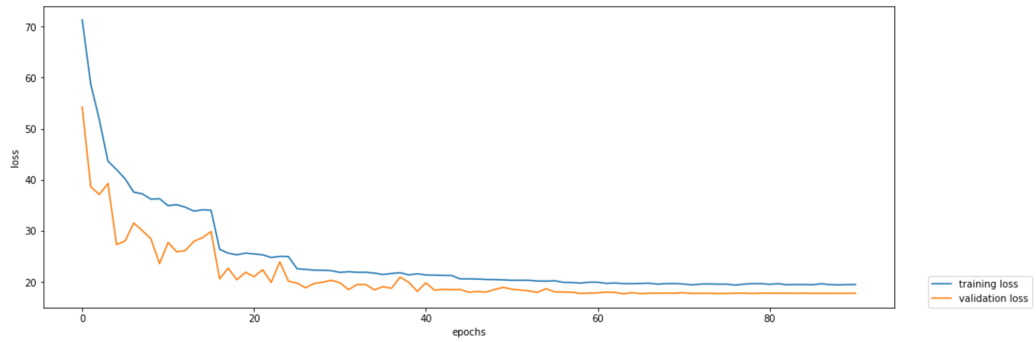
3000 s

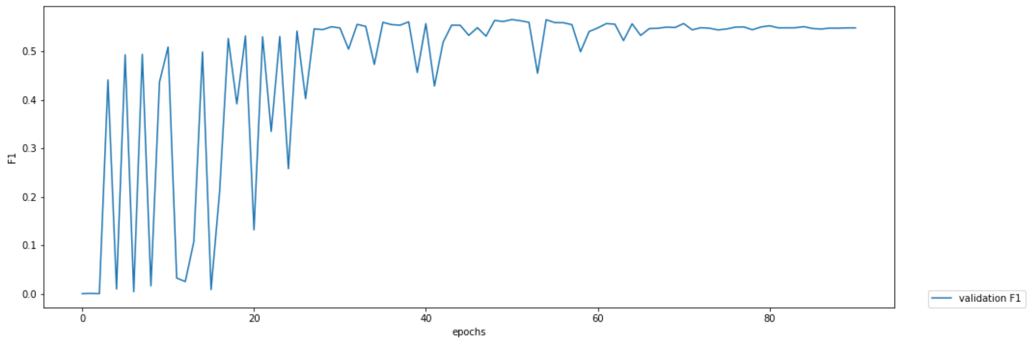
EPOCH 91 done: loss 19.5252 - lr 0.0001953
2021-06-17 12:28:24,109 DEV : loss 17.82100486755371 - score 0.5483

KEY tp: 3024 - fp: 2501 - fn: 1820 - precision: 0.5473 - recall: 0.6243 - f1-score: 0.5833

KEY tp: 2819 - fp: 2586 - fn: 1747 - precision: 0.5216 - recall: 0.6174 - f1-score: 0.5654
tensor(18.5803, device='cuda:0')

KEY tp: 6440 - fp: 5168 - fn: 3341 - precision: 0.5548 - recall: 0.6584 - f1-score: 0.6022
tensor(18.9307, device='cuda:0')





_ BERT 2x256

embedding = 'Bert'
embedding_path = ''
dataset_base_path = ''
dataset = 'Inspec'
output_base_path = '../result/'
iteration = ''
gpu = 1
lr = 0.05
anneal_factor = 0.5
patience = 4
batch_size = 4
num_epochs = 100
threads = 12
param_selection_mode = False
use_tensorboard = False
no_dev = False
use_crf = True
rnn_layers = 2
hidden_size = 256
dropout = 0.0
word_dropout = 0.05
locked_dropout = 0.5
downsample_train = 0.0
not_in_memory = False

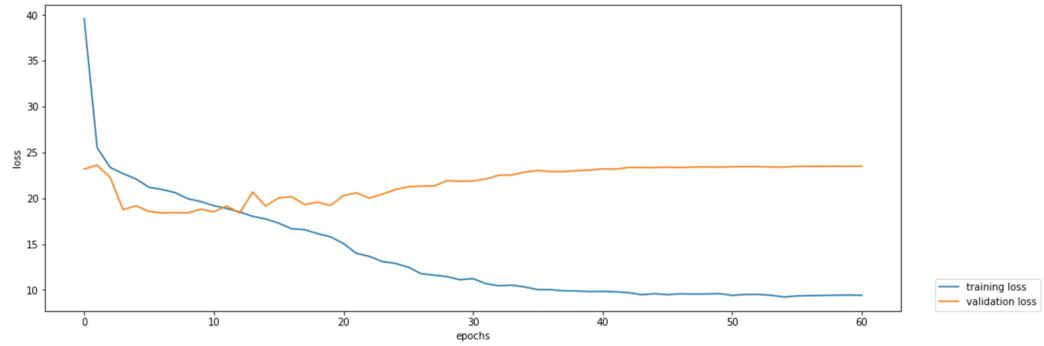
9800 s

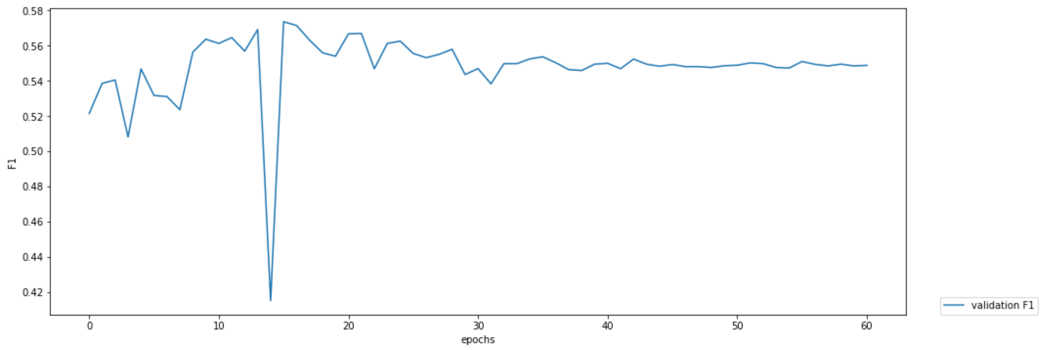
EPOCH 61 done: loss 9.4253 - lr 0.0001953
2021-06-18 14:36:53,953 DEV : loss 23.496610641479492 - score 0.5488

KEY tp: 3335 - fp: 2930 - fn: 1509 - precision: 0.5323 - recall: 0.6885 - f1-score: 0.6004

KEY tp: 3049 - fp: 3016 - fn: 1517 - precision: 0.5027 - recall: 0.6678 - f1-score: 0.5736
tensor(19.9994, device='cuda:0')

KEY tp: 7793 - fp: 5054 - fn: 1985 - precision: 0.6066 - recall: 0.7970 - f1-score: 0.6889
tensor(15.2339, device='cuda:0')





BERT 2x256 LR 0.2

embedding = 'Bert'
embedding_path = ''
dataset_base_path = ''
dataset = 'Inspec'
output_base_path = '../result/'
iteration = ''
gpu = 1
lr = 0.2
anneal_factor = 0.5
patience = 4
batch_size = 4
num_epochs = 100
threads = 12
param_selection_mode = False
use_tensorboard = False
no_dev = False
use_crf = True
rnn_layers = 2
hidden_size = 256
dropout = 0.0
word_dropout = 0.05
locked_dropout = 0.5
downsample_train = 0.0
not_in_memory = False

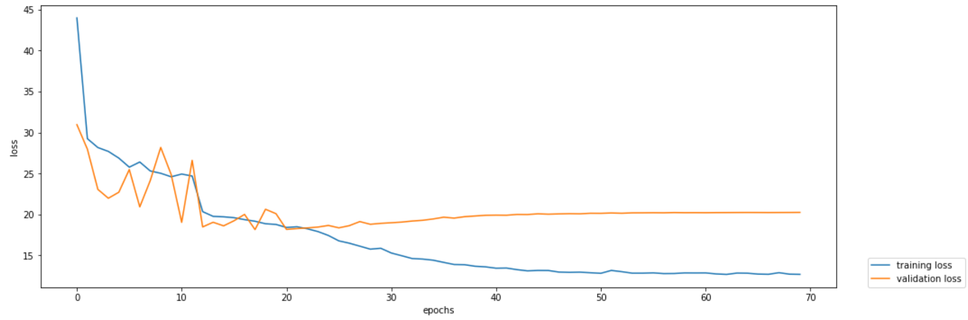
6200 s

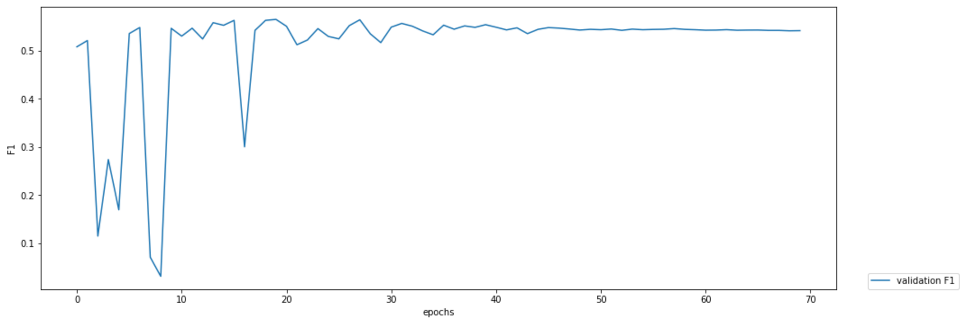
EPOCH 70 done: loss 12.6877 - lr 0.0001953
2021-06-21 12:51:31,683 DEV : loss 20.251916885375977 - score 0.5416

KEY tp: 3389 - fp: 2973 - fn: 1455 - precision: 0.5327 - recall: 0.6996 - f1-score: 0.6049

KEY tp: 3041 - fp: 3159 - fn: 1525 - precision: 0.4905 - recall: 0.6660 - f1-score: 0.5649
tensor(20.0703, device='cuda:0')

KEY tp: 7392 - fp: 5876 - fn: 2386 - precision: 0.5571 - recall: 0.7560 - f1-score: 0.6415
tensor(17.5604, device='cuda:0')





X BERT 4x128 DO 0.3

embedding = 'Bert'
embedding_path = ''
dataset_base_path = ''
dataset = 'Inspec'
output_base_path = '../result/'
iteration = ''
gpu = 1
lr = 0.05
anneal_factor = 0.5
patience = 4
batch_size = 4
num_epochs = 100
threads = 12
param_selection_mode = False
use_tensorboard = False
no_dev = False
use_crf = True
rnn_layers = 4
hidden_size = 128
dropout = 0.3
word_dropout = 0.05
locked_dropout = 0.5
downsample_train = 0.0
not_in_memory = False

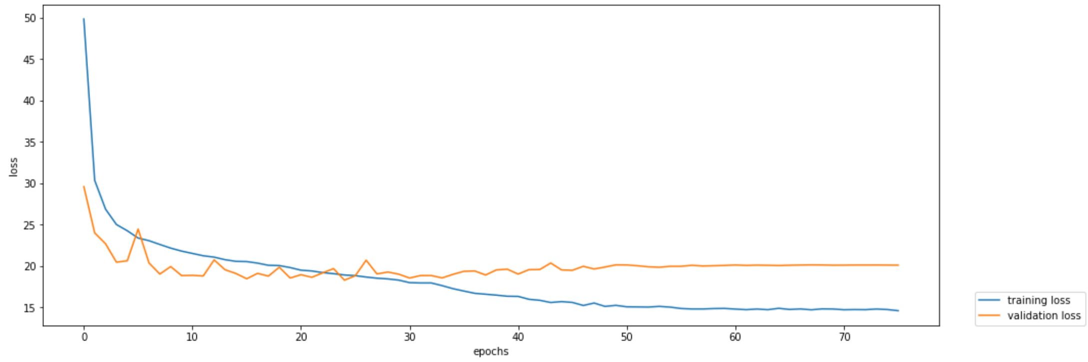
7000 s

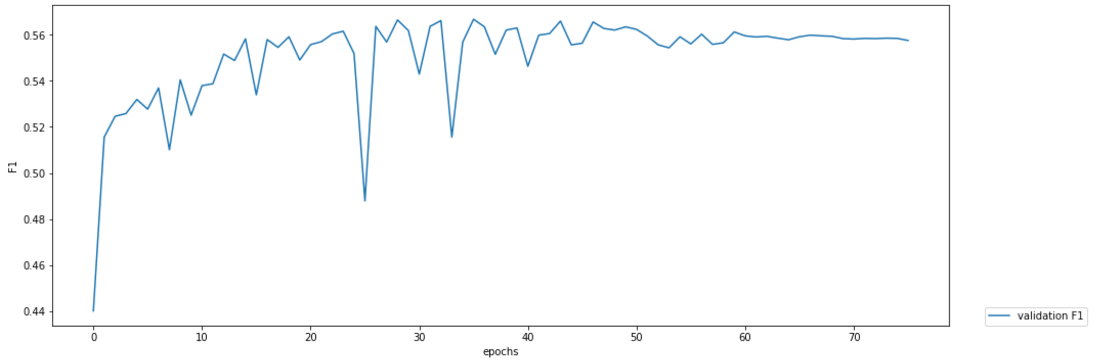
EPOCH 76 done: loss 14.5995 - lr 0.0001953
2021-06-18 21:30:08,496 DEV : loss 20.108718872070312 - score 0.5575

KEY tp: 3166 - fp: 2612 - fn: 1678 - precision: 0.5479 - recall: 0.6536 - f1-score: 0.5961

KEY tp: 2864 - fp: 2678 - fn: 1702 - precision: 0.5168 - recall: 0.6272 - f1-score: 0.5667
tensor(19.3232, device='cuda:0')

KEY tp: 7645 - fp: 4136 - fn: 2133 - precision: 0.6489 - recall: 0.7819 - f1-score: 0.7092
tensor(13.8824, device='cuda:0')





X BERT 3x128 DO 0.3 LR 0.2

embedding = 'Bert'
embedding_path = ''
dataset_base_path = ''
dataset = 'Inspec'
output_base_path = '../result/'
iteration = ''
gpu = 1
lr = 0.2
anneal_factor = 0.5
patience = 4
batch_size = 4
num_epochs = 100
threads = 12
param_selection_mode = False
use_tensorboard = False
no_dev = False
use_crf = True
rnn_layers = 3
hidden_size = 128
dropout = 0.3
word_dropout = 0.05
locked_dropout = 0.5
downsample_train = 0.0
not_in_memory = False

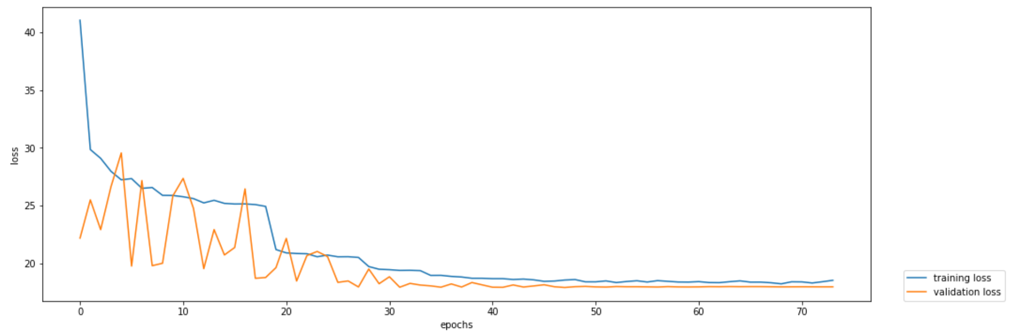
6500 s

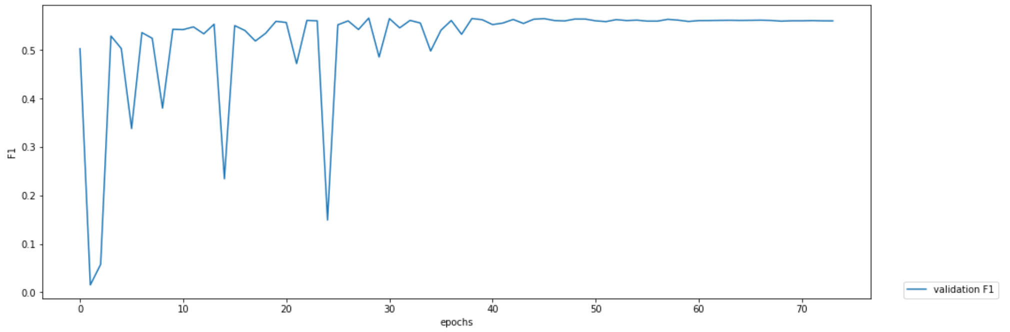
EPOCH 74 done: loss 18.5210 - lr 0.0001953
2021-06-21 12:53:42,190 DEV : loss 17.958280563354492 - score 0.5604

KEY tp: 3335 - fp: 2865 - fn: 1509 - precision: 0.5379 - recall: 0.6885 - f1-score: 0.6039

KEY tp: 3010 - fp: 3056 - fn: 1556 - precision: 0.4962 - recall: 0.6592 - f1-score: 0.5662
tensor(19.4615, device='cuda:0')

KEY tp: 6983 - fp: 6108 - fn: 2795 - precision: 0.5334 - recall: 0.7142 - f1-score: 0.6107
tensor(19.0372, device='cuda:0')





BERT 4x128 DO 0.3 LR 0.1

embedding = 'Bert'
embedding_path = ''
dataset_base_path = ''
dataset = 'Inspec'
output_base_path = '../result/'
iteration = ''
gpu = 1
lr = 0.1
anneal_factor = 0.5
patience = 4
batch_size = 16
num_epochs = 150
threads = 12
param_selection_mode = False
use_tensorboard = False
no_dev = False
use_crf = True
rnn_layers = 4
hidden_size = 128
dropout = 0.3
word_dropout = 0.05
locked_dropout = 0.5
downsample_train = 0.0
not_in_memory = False

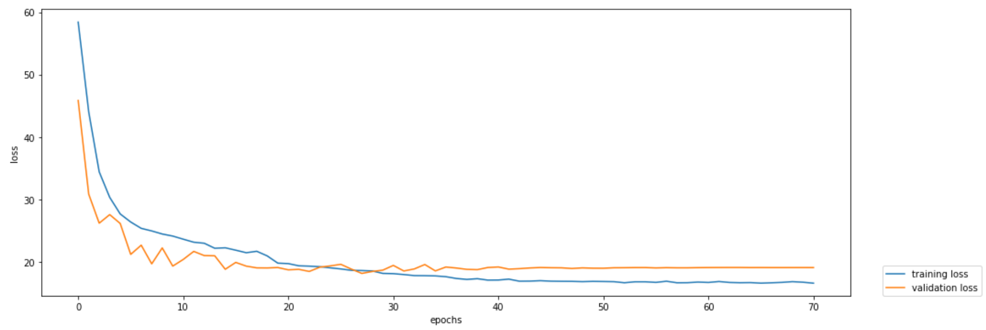
2800 s

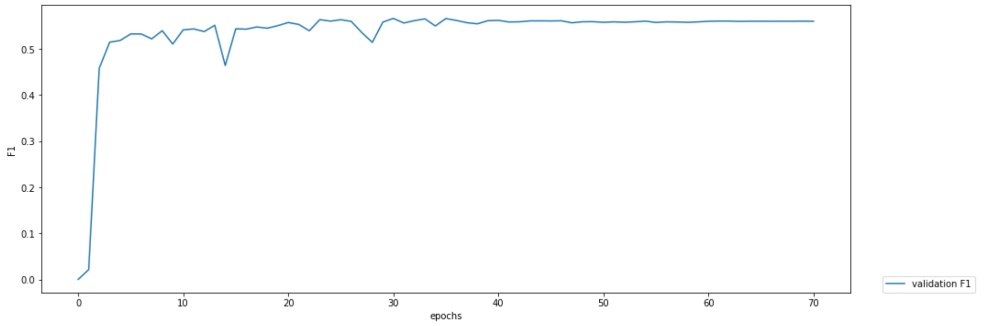
EPOCH 71 done: loss 16.6242 - lr 0.0001953
2021-06-21 15:24:05,878 DEV : loss 19.123266220092773 - score 0.5599

KEY tp: 3213 - fp: 2782 - fn: 1631 - precision: 0.5359 - recall: 0.6633 - f1-score: 0.5929

KEY tp: 2945 - fp: 2892 - fn: 1621 - precision: 0.5045 - recall: 0.6450 - f1-score: 0.5662
tensor(19.4410, device='cuda:0')

KEY tp: 7346 - fp: 5075 - fn: 2432 - precision: 0.5914 - recall: 0.7513 - f1-score: 0.6618
tensor(16.2648, device='cuda:0')





_ BERT 2x64

embedding = 'Bert'
  embedding_path = ''
  dataset_base_path = ''
  dataset = 'Inspec'
  output_base_path = '../result/'
  iteration = ''
  gpu = 1
  lr = 0.05
  anneal_factor = 0.5
  patience = 4
  batch_size = 4
  num_epochs = 100
  threads = 12
  param_selection_mode = False
  use_tensorboard = False
  no_dev = False
  use_crf = True
  rnn_layers = 2
  hidden_size = 64
  dropout = 0.0
  word_dropout = 0.05
  locked_dropout = 0.5
  downsample_train = 0.0
  not_in_memory = False

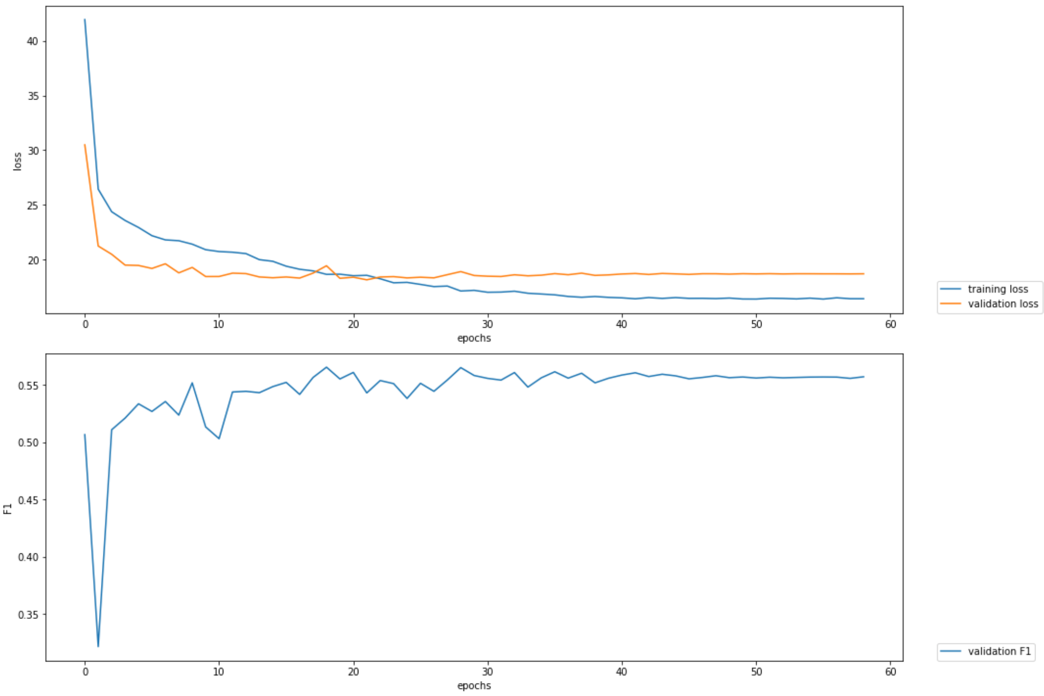
4800 s

EPOCH 59 done: loss 16.4335 - lr 0.0001953
2021-06-21 15:56:36,544 DEV : loss 18.715980529785156 - score 0.5573

KEY tp: 3252 - fp: 2764 - fn: 1592 - precision: 0.5406 - recall: 0.6713 - f1-score: 0.5989

KEY tp: 2944 - fp: 2898 - fn: 1622 - precision: 0.5039 - recall: 0.6448 - f1-score: 0.5657
tensor(19.4264, device='cuda:0')

KEY tp: 7197 - fp: 5292 - fn: 2581 - precision: 0.5763 - recall: 0.7360 - f1-score: 0.6464
tensor(17.3252, device='cuda:0')



_ BERT 4x64

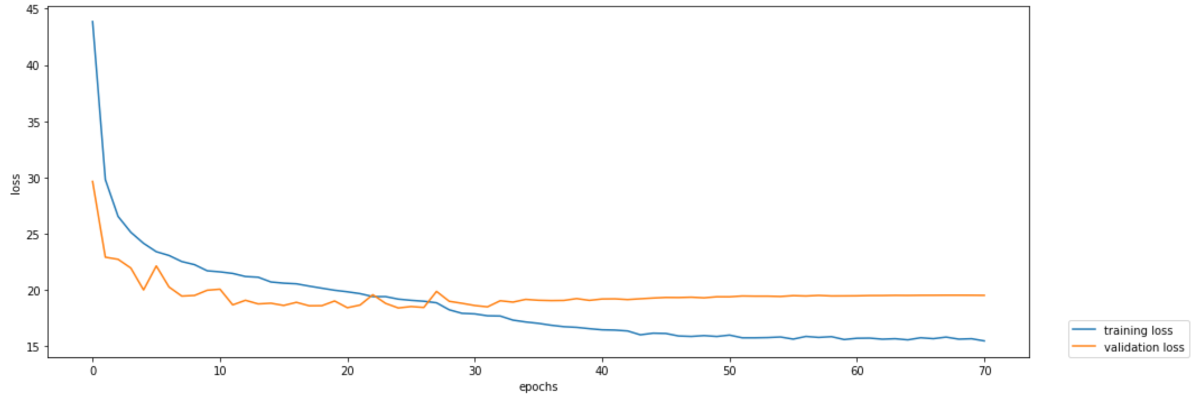
embedding = 'Bert'
embedding_path = ''
dataset_base_path = ''
dataset = 'Inspec'
output_base_path = '../result/'
iteration = ''
gpu = 1
lr = 0.05
anneal_factor = 0.5
patience = 4
batch_size = 4
num_epochs = 100
threads = 12
param_selection_mode = False
use_tensorboard = False
no_dev = False
use_crf = True
rnn_layers = 4
hidden_size = 64
dropout = 0.0
word_dropout = 0.05
locked_dropout = 0.5
downsample_train = 0.0
not_in_memory = False

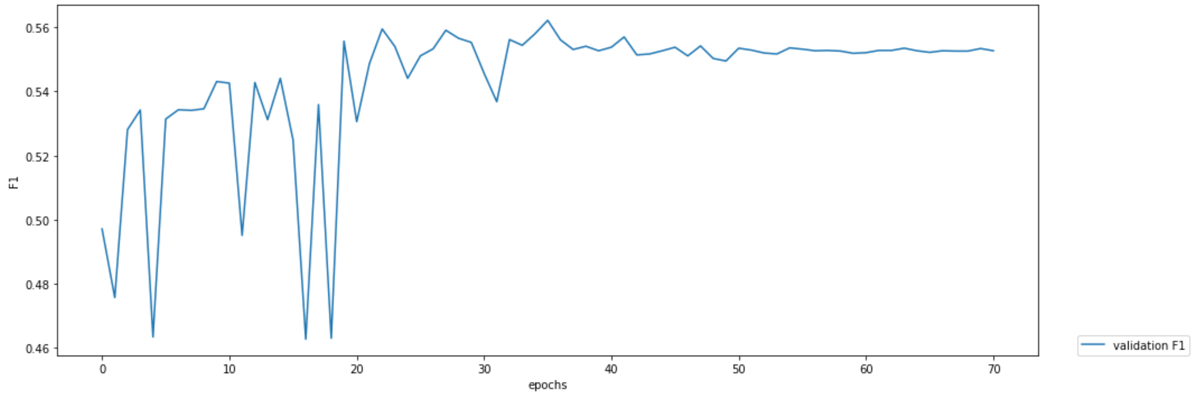
6500 s

KEY tp: 3039 - fp: 2433 - fn: 1805 - precision: 0.5554 - recall: 0.6274 - f1-score: 0.5892

KEY tp: 2756 - fp: 2483 - fn: 1810 - precision: 0.5261 - recall: 0.6036 - f1-score: 0.5622
tensor(19.0425, device='cuda:0')

KEY tp: 7358 - fp: 3982 - fn: 2420 - precision: 0.6489 - recall: 0.7525 - f1-score: 0.6968
tensor(14.3142, device='cuda:0')





_ BERT 3x128

embedding = 'Bert'
embedding_path = ''
dataset_base_path = ''
dataset = 'Inspec'
output_base_path = '../result/'
iteration = ''
gpu = 1
lr = 0.05
anneal_factor = 0.5
patience = 4
batch_size = 4
num_epochs = 100
threads = 12
param_selection_mode = False
use_tensorboard = False
no_dev = False
use_crf = True
rnn_layers = 3
hidden_size = 128
dropout = 0.0
word_dropout = 0.05
locked_dropout = 0.5
downsample_train = 0.0
not_in_memory = False

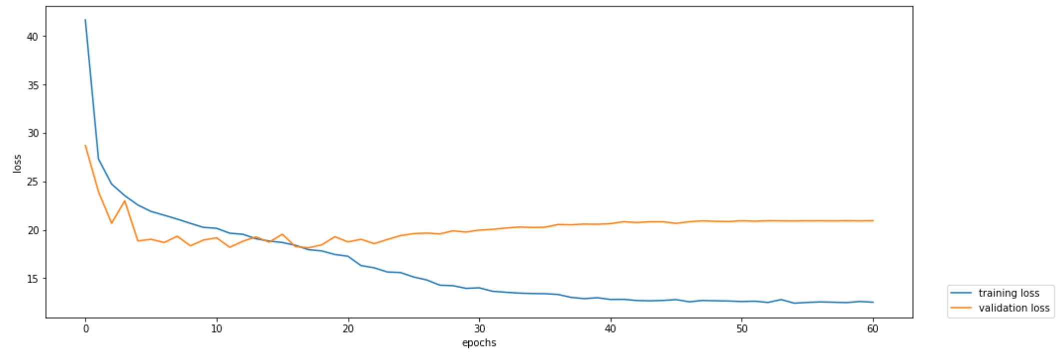
5500 s

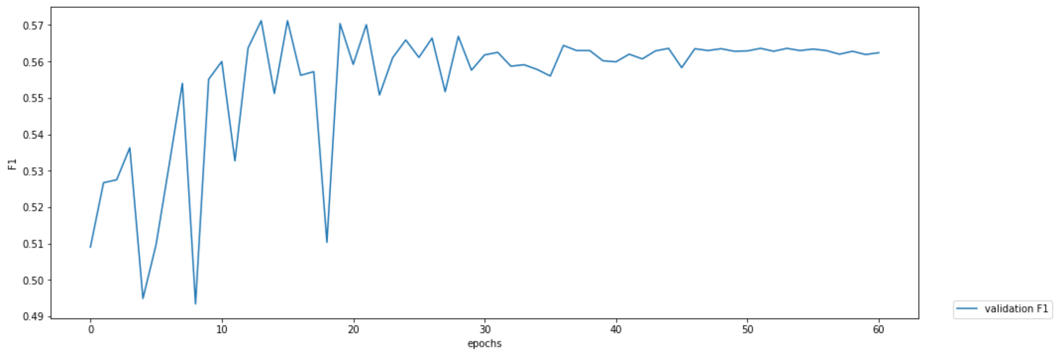
EPOCH 61 done: loss 12.5200 - lr 0.0001953
2021-06-21 16:59:52,007 DEV : loss 20.942089080810547 - score 0.5624

KEY tp: 3282 - fp: 2789 - fn: 1562 - precision: 0.5406 - recall: 0.6775 - f1-score: 0.6014

KEY tp: 2992 - fp: 2918 - fn: 1574 - precision: 0.5063 - recall: 0.6553 - f1-score: 0.5712
tensor(19.4974, device='cuda:0')

KEY tp: 7274 - fp: 5393 - fn: 2504 - precision: 0.5742 - recall: 0.7439 - f1-score: 0.6482
tensor(16.8266, device='cuda:0')





_ ROBERTA 2x64

embedding = 'RoBERTa'
embedding_path = ''
dataset_base_path = ''
dataset = 'Inspec'
output_base_path = '../result/'
iteration = ''
gpu = 1
lr = 0.05
anneal_factor = 0.5
patience = 4
batch_size = 4
num_epochs = 100
threads = 12
param_selection_mode = False
use_tensorboard = False
no_dev = False
use_crf = True
rnn_layers = 2
hidden_size = 64
dropout = 0.0
word_dropout = 0.05
locked_dropout = 0.5
downsample_train = 0.0
not_in_memory = False

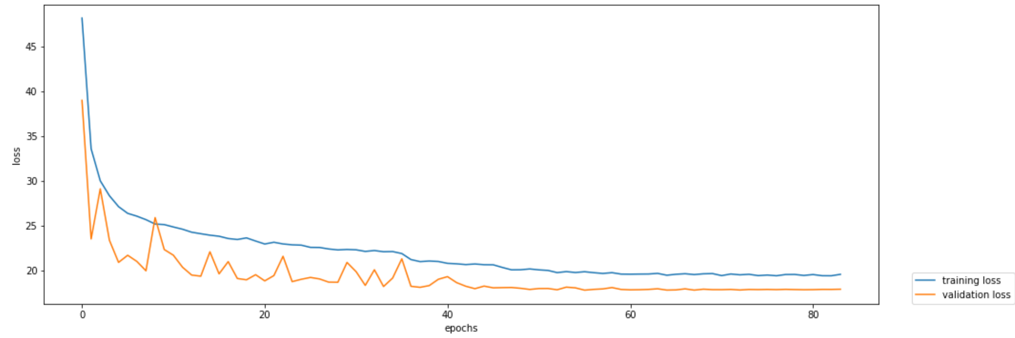
6700 s

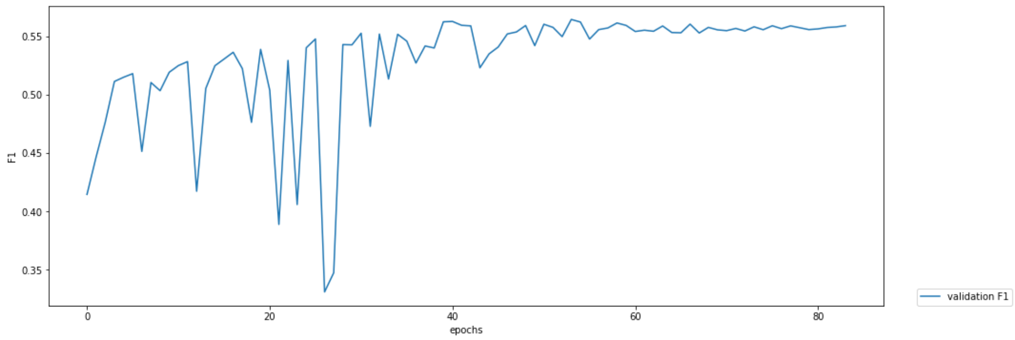
EPOCH 84 done: loss 19.5465 - lr 0.0001953
2021-06-21 18:20:21,881 DEV : loss 17.884660720825195 - score 0.559

KEY tp: 2901 - fp: 2197 - fn: 1943 - precision: 0.5690 - recall: 0.5989 - f1-score: 0.5836

KEY tp: 2696 - fp: 2294 - fn: 1870 - precision: 0.5403 - recall: 0.5905 - f1-score: 0.5643
tensor(18.0621, device='cuda:0')

KEY tp: 6272 - fp: 4309 - fn: 3509 - precision: 0.5928 - recall: 0.6412 - f1-score: 0.6160
tensor(17.9286, device='cuda:0')





_ ROBERTA 4x64

embedding = 'RoBERTa'
embedding_path = ''
dataset_base_path = ''
dataset = 'Inspec'
output_base_path = '../result/'
iteration = ''
gpu = 1
lr = 0.05
anneal_factor = 0.5
patience = 4
batch_size = 4
num_epochs = 100
threads = 12
param_selection_mode = False
use_tensorboard = False
no_dev = False
use_crf = True
rnn_layers = 4
hidden_size = 64
dropout = 0.0
word_dropout = 0.05
locked_dropout = 0.5
downsample_train = 0.0
not_in_memory = False

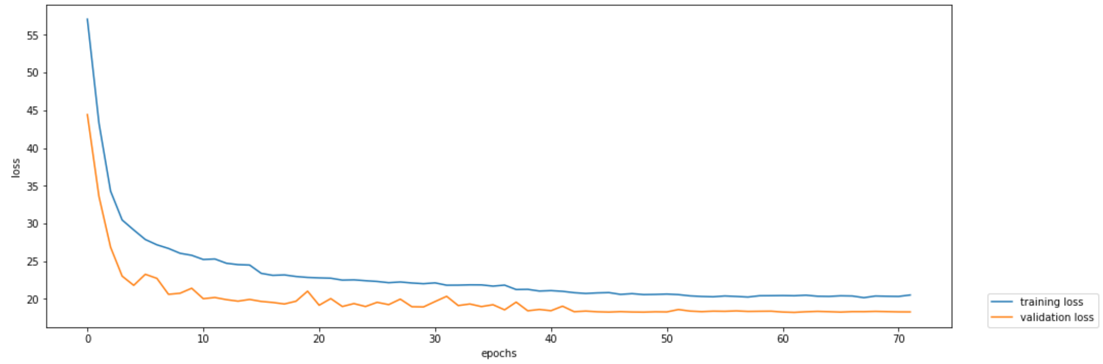
6400 s

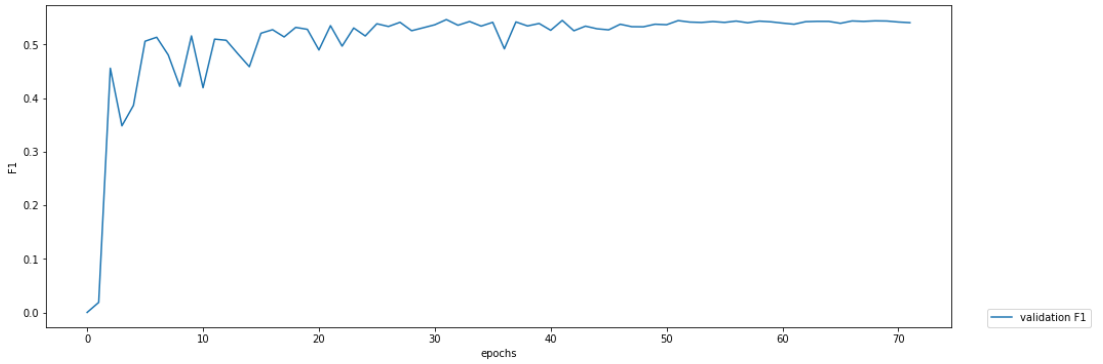
EPOCH 72 done: loss 20.5242 - lr 0.0001953
2021-06-21 18:02:22,590 DEV : loss 18.279985427856445 - score 0.5406

KEY tp: 3281 - fp: 3080 - fn: 1563 - precision: 0.5158 - recall: 0.6773 - f1-score: 0.5856

KEY tp: 2978 - fp: 3357 - fn: 1588 - precision: 0.4701 - recall: 0.6522 - f1-score: 0.5464
tensor(20.2948, device='cuda:0')

KEY tp: 6630 - fp: 6790 - fn: 3151 - precision: 0.4940 - recall: 0.6778 - f1-score: 0.5715
tensor(21.1824, device='cuda:0')





X BERT 3x128 DO 0.3 LR 0.05 OP Adam

embedding = 'Bert'
  embedding_path = ''
  dataset_base_path = ''
  dataset = 'Inspec'
  output_base_path = '../result/'
  iteration = ''
  gpu = 1
  lr = 0.05
  anneal_factor = 0.5
  patience = 4
  batch_size = 8
  num_epochs = 100
  threads = 12
  param_selection_mode = False
  use_tensorboard = False
  no_dev = False
  use_crf = True
  rnn_layers = 3
  hidden_size = 128
  dropout = 0.3
  word_dropout = 0.05
  locked_dropout = 0.5
  downsample_train = 0.0
  not_in_memory = False

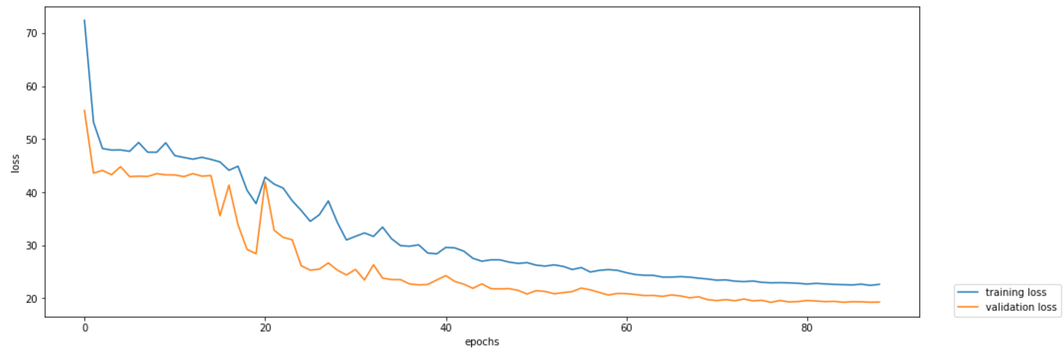
5000 s

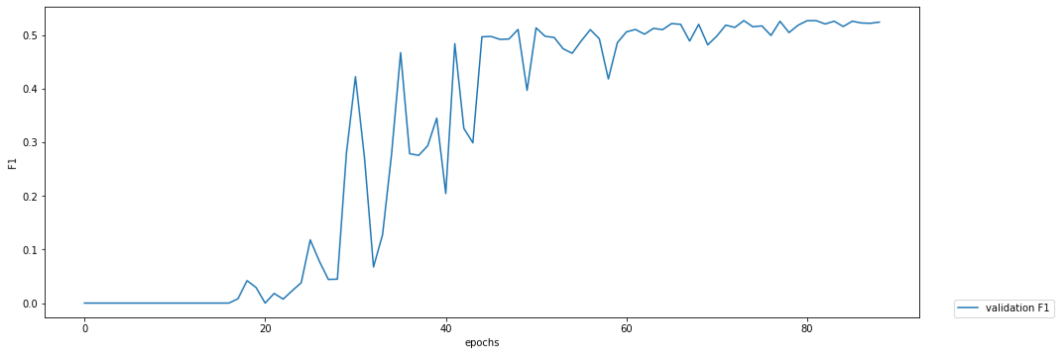
EPOCH 89 done: loss 22.6529 - lr 0.0001953
2021-06-22 10:58:27,876 DEV : loss 19.303680419921875 - score 0.5242

KEY   tp: 2815 - fp: 2446 - fn: 2029 - precision: 0.5351 - recall: 0.5811 - f1-score: 0.5571

KEY   tp: 2602 - fp: 2707 - fn: 1964 - precision: 0.4901 - recall: 0.5699 - f1-score: 0.5270
tensor(19.8293, device='cuda:0')

KEY   tp: 5502 - fp: 5554 - fn: 4276 - precision: 0.4976 - recall: 0.5627 - f1-score: 0.5282
tensor(21.2727, device='cuda:0')





X BERT 3x128 DO 0.3 LR 0.2 OP Adam

embedding = 'Bert'
embedding_path = ''
dataset_base_path = ''
dataset = 'Inspec'
output_base_path = '../result/'
iteration = ''
gpu = 1
lr = 0.2
anneal_factor = 0.5
patience = 4
batch_size = 8
num_epochs = 150
threads = 12
param_selection_mode = False
use_tensorboard = False
no_dev = False
use_crf = True
rnn_layers = 3
hidden_size = 128
dropout = 0.3
word_dropout = 0.05
locked_dropout = 0.5
downsample_train = 0.0
not_in_memory = False

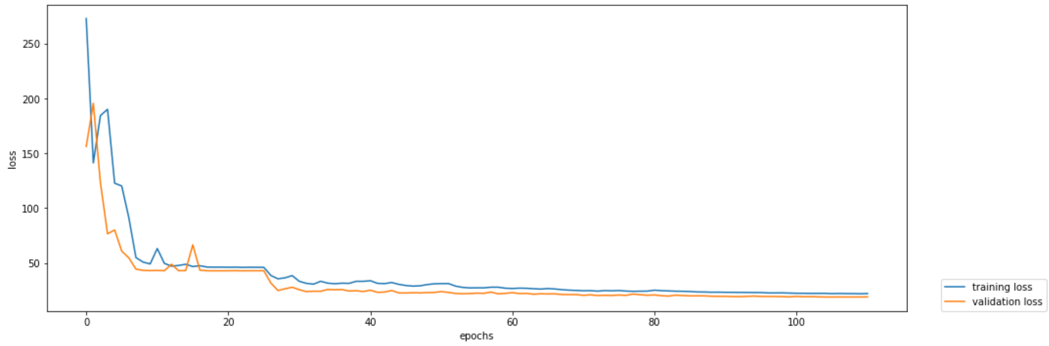
6100 s

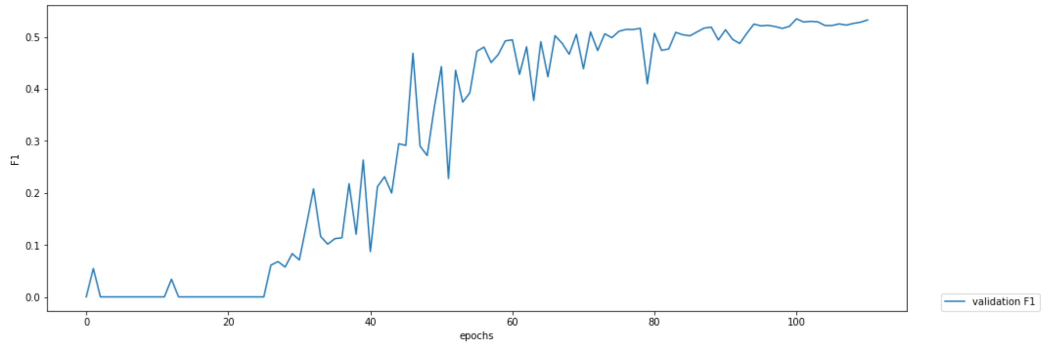
EPOCH 111 done: loss 22.1569 - lr 0.0001953
2021-06-22 12:52:56,456 DEV : loss 19.085237503051758 - score 0.5326

KEY tp: 2856 - fp: 2400 - fn: 1988 - precision: 0.5434 - recall: 0.5896 - f1-score: 0.5655

EY tp: 2626 - fp: 2629 - fn: 1940 - precision: 0.4997 - recall: 0.5751 - f1-score: 0.5348
tensor(19.4417, device='cuda:0')

EY tp: 5651 - fp: 5338 - fn: 4127 - precision: 0.5142 - recall: 0.5779 - f1-score: 0.5442
tensor(20.6624, device='cuda:0')





X BERT 2x128 LR 0.05 OP AdamW

embedding = 'Bert'
embedding_path = ''
dataset_base_path = ''
dataset = 'Inspec'
output_base_path = '../result/'
iteration = ''
gpu = 1
lr = 0.05
anneal_factor = 0.5
patience = 4
batch_size = 8
num_epochs = 100
threads = 12
param_selection_mode = False
use_tensorboard = False
no_dev = False
use_crf = True
rnn_layers = 2
hidden_size = 128
dropout = 0.0
word_dropout = 0.05
locked_dropout = 0.5
downsample_train = 0.0
not_in_memory = False

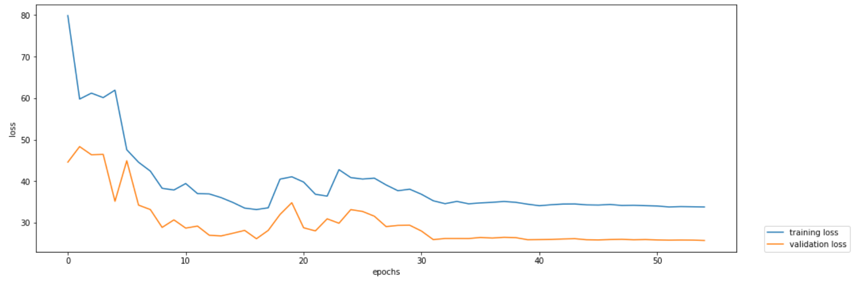
3000 s

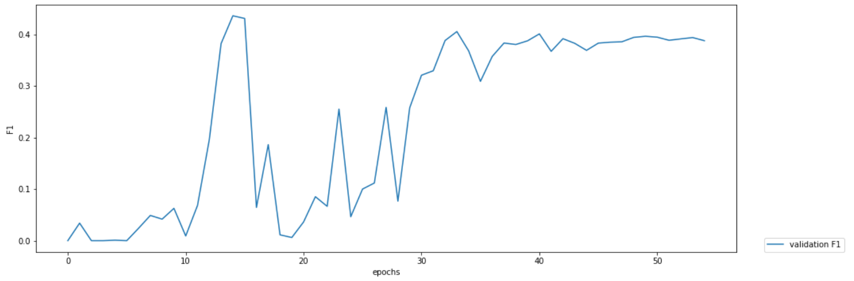
EPOCH 55 done: loss 33.7824 - lr 0.0001953
2021-06-22 20:10:20,644 DEV : loss 25.704378128051758 - score 0.3875

KEY tp: 2443 - fp: 3159 - fn: 2401 - precision: 0.4361 - recall: 0.5043 - f1-score: 0.4677

EY tp: 2208 - fp: 3359 - fn: 2358 - precision: 0.3966 - recall: 0.4836 - f1-score: 0.4358
tensor(27.3892, device='cuda:0')

EY tp: 4762 - fp: 7135 - fn: 5016 - precision: 0.4003 - recall: 0.4870 - f1-score: 0.4394
tensor(29.1305, device='cuda:0')





3x128 DO 0.3 0.0489 51 epochs 6000 s

X BER 2x128 LR 0.05 OP Adagrad

embedding = 'Bert'
embedding_path = ''
dataset_base_path = ''
dataset = 'Inspec'
output_base_path = '../result/'
iteration = ''
gpu = 1
lr = 0.05
anneal_factor = 0.5
patience = 4
batch_size = 8
num_epochs = 100
threads = 12
param_selection_mode = False
use_tensorboard = False
no_dev = False
use_crf = True
rnn_layers = 2
hidden_size = 128
dropout = 0.0
word_dropout = 0.05
locked_dropout = 0.5
downsample_train = 0.0
not_in_memory = False

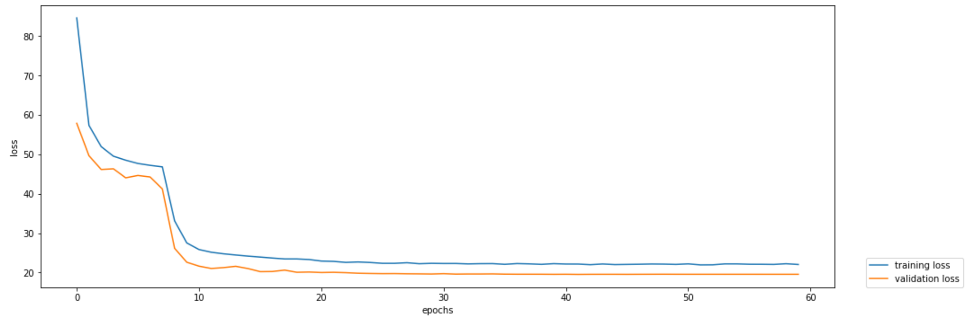
3200 s

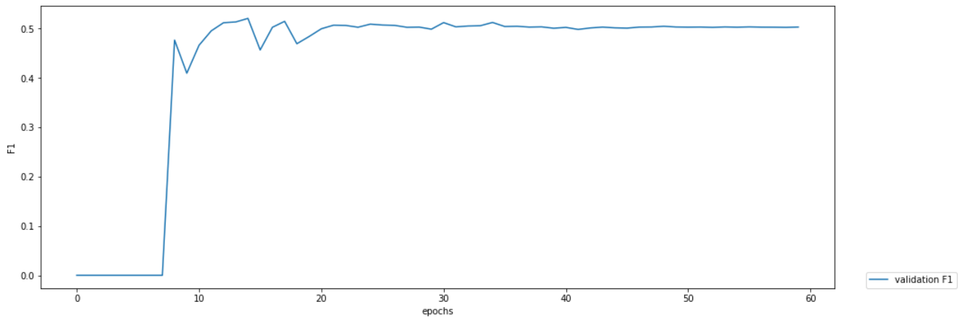
EPOCH 60 done: loss 22.0195 - lr 0.0001953
2021-06-22 20:06:34,939 DEV : loss 19.52480697631836 - score 0.503

KEY tp: 2828 - fp: 2579 - fn: 2016 - precision: 0.5230 - recall: 0.5838 - f1-score: 0.5518

EY tp: 2591 - fp: 2799 - fn: 1975 - precision: 0.4807 - recall: 0.5675 - f1-score: 0.5205
tensor(20.9450, device='cuda:0')

EY tp: 5490 - fp: 5740 - fn: 4288 - precision: 0.4889 - recall: 0.5615 - f1-score: 0.5227
tensor(22.2837, device='cuda:0')





X ! BERT 3x128 DO 0.3 OP Adadelta

embedding = 'Bert'
embedding_path = ''
dataset_base_path = ''
dataset = 'Inspec'
output_base_path = '../result/'
iteration = ''
gpu = 1
lr = 0.005
anneal_factor = 0.5
patience = 4
batch_size = 8
num_epochs = 150
threads = 12
param_selection_mode = False
use_tensorboard = False
no_dev = False
use_crf = True
rnn_layers = 3
hidden_size = 128
dropout = 0.3
word_dropout = 0.05
locked_dropout = 0.5
downsample_train = 0.0
not_in_memory = False

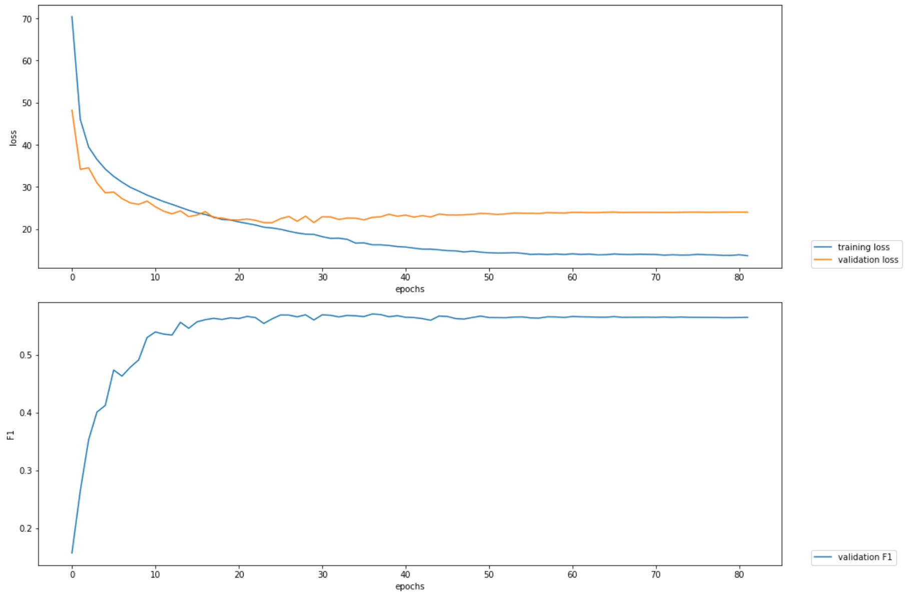
4600 s

EPOCH 82 done: loss 13.7012 - lr 0.0001953
2021-06-22 19:27:09,141 DEV : loss 24.043468475341797 - score 0.5652

KEY tp: 3269 - fp: 2885 - fn: 1575 - precision: 0.5312 - recall: 0.6749 - f1-score: 0.5945

KEY tp: 2987 - fp: 2908 - fn: 1579 - precision: 0.5067 - recall: 0.6542 - f1-score: 0.5711
tensor(22.7880, device='cuda:0')

KEY tp: 8208 - fp: 3816 - fn: 1570 - precision: 0.6826 - recall: 0.8394 - f1-score: 0.7530
tensor(12.8425, device='cuda:0')



X BER 2x128 OP RMSprop

0.1783 LR 0.3 0.0847 LR 0.05

X BERT 3x128 LR 0.05 Adagrad

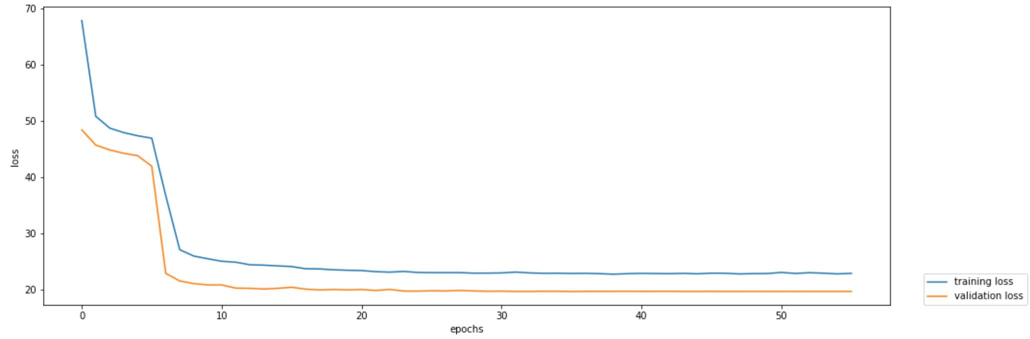
embedding = 'Bert'
embedding_path = ''
dataset_base_path = ''
dataset = 'Inspec'
output_base_path = '../result/'
iteration = ''
gpu = 1
lr = 0.05
anneal_factor = 0.5
patience = 4
batch_size = 8
num_epochs = 100
threads = 12
param_selection_mode = False
use_tensorboard = False
no_dev = False
use_crf = True
rnn_layers = 3
hidden_size = 128
dropout = 0.3
word_dropout = 0.05
locked_dropout = 0.5
downsample_train = 0.0
not_in_memory = False

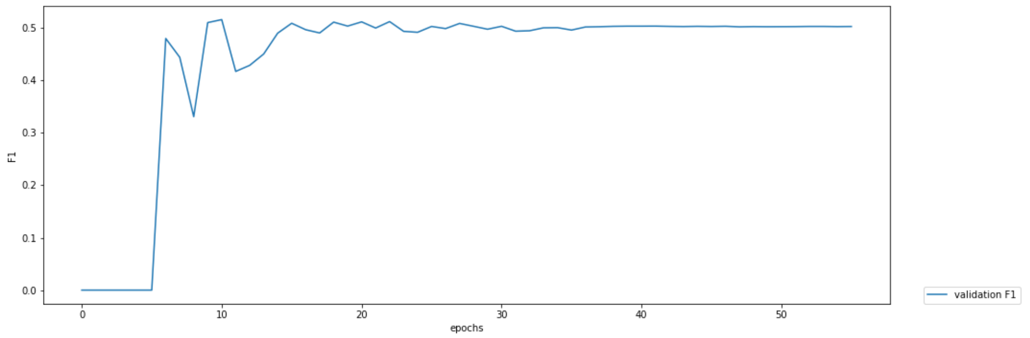
3100 s

EPOCH 56 done: loss 22.9055 - lr 0.0001953
2021-06-23 10:02:55,955 DEV : loss 19.698169708251953 - score 0.503

KEY tp: 2508 - fp: 2644 - fn: 2058 - precision: 0.4868 - recall: 0.5493 - f1-score: 0.5162
tensor(20.8123, device='cuda:0')

KEY tp: 5296 - fp: 5575 - fn: 4482 - precision: 0.4872 - recall: 0.5416 - f1-score: 0.5130
tensor(22.3577, device='cuda:0')





X ! BERT 2x128 OP Adadelta

embedding = 'Bert'
embedding_path = ''
dataset_base_path = ''
dataset = 'Inspec'
output_base_path = '../result/'
iteration = ''
gpu = 1
lr = 0.005
anneal_factor = 0.5
patience = 4
batch_size = 8
num_epochs = 150
threads = 12
param_selection_mode = False
use_tensorboard = False
no_dev = False
use_crf = True
rnn_layers = 2
hidden_size = 128
dropout = 0.0
word_dropout = 0.05
locked_dropout = 0.5
downsample_train = 0.0
not_in_memory = False

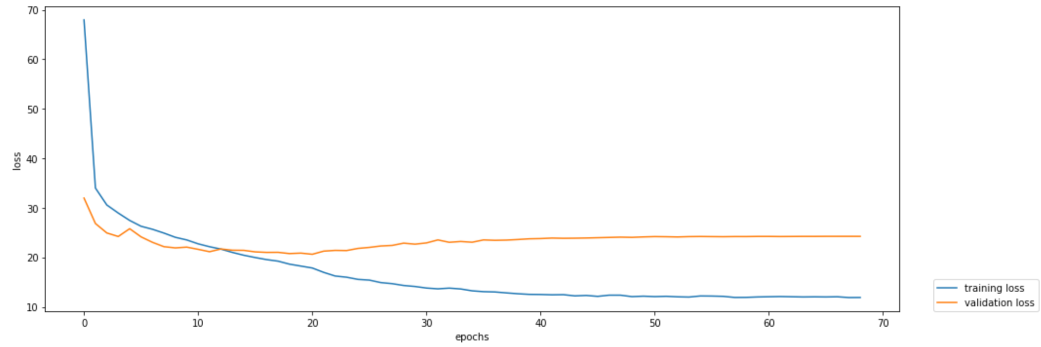
3700 s

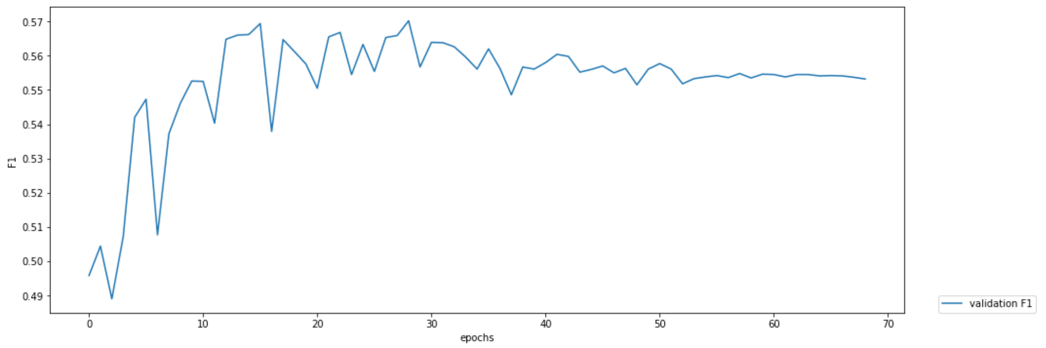
EPOCH 69 done: loss 11.9242 - lr 0.0001953
2021-06-23 10:11:41,314 DEV : loss 24.26744842529297 - score 0.5532

KEY tp: 3009 - fp: 2369 - fn: 1835 - precision: 0.5595 - recall: 0.6212 - f1-score: 0.5887

KEY tp: 2790 - fp: 2430 - fn: 1776 - precision: 0.5345 - recall: 0.6110 - f1-score: 0.5702
tensor(22.8585, device='cuda:0')

KEY tp: 8295 - fp: 2412 - fn: 1483 - precision: 0.7747 - recall: 0.8483 - f1-score: 0.8099
tensor(9.8417, device='cuda:0')





X BERT 3x128 DO 0.3 NO CRF

embedding = 'Bert'
embedding_path = ''
dataset_base_path = ''
dataset = 'Inspec'
output_base_path = '../result/'
iteration = ''
gpu = 1
lr = 0.05
anneal_factor = 0.5
patience = 4
batch_size = 4
num_epochs = 100
threads = 12
param_selection_mode = False
use_tensorboard = False
no_dev = False
use_crf = False
rnn_layers = 3
hidden_size = 128
dropout = 0.3
word_dropout = 0.05
locked_dropout = 0.5
downsample_train = 0.0
not_in_memory = False

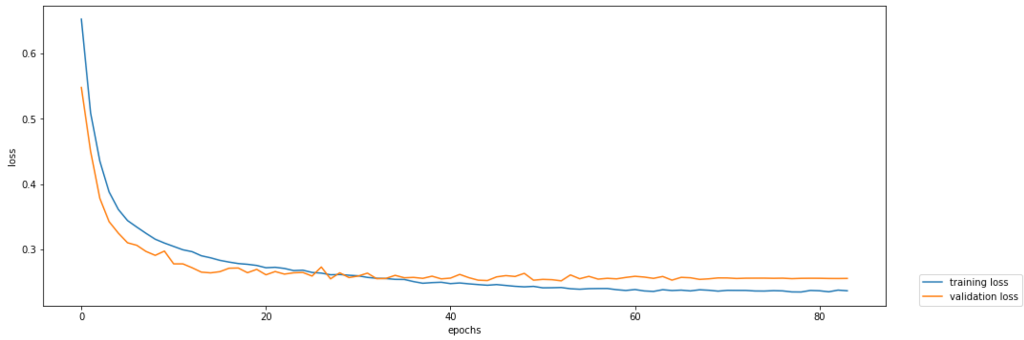
2200 s

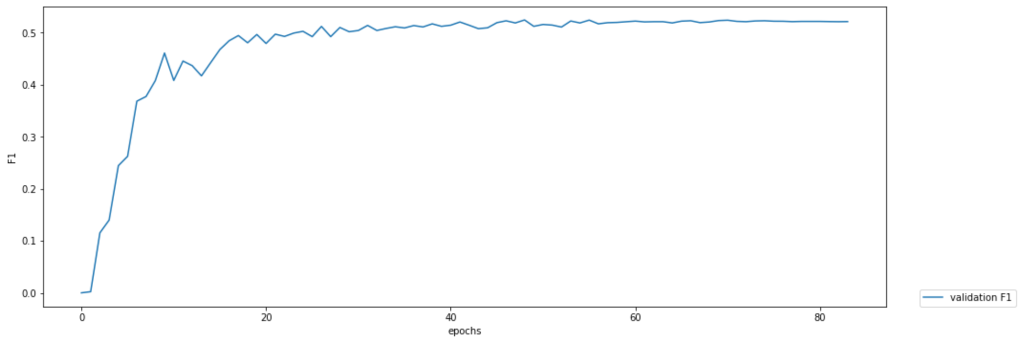
EPOCH 84 done: loss 0.2362 - lr 0.0001953
2021-06-23 16:52:41,686 DEV : loss 0.25519272685050964 - score 0.5213

KEY tp: 3154 - fp: 3462 - fn: 1690 - precision: 0.4767 - recall: 0.6511 - f1-score: 0.5504

KEY tp: 2917 - fp: 3643 - fn: 1649 - precision: 0.4447 - recall: 0.6389 - f1-score: 0.5244
tensor(0.2633, device='cuda:0')

KEY tp: 6570 - fp: 7292 - fn: 3208 - precision: 0.4740 - recall: 0.6719 - f1-score: 0.5558
tensor(0.2283, device='cuda:0')





Mejor modelo

embedding = 'Bert'
embedding_path = ''
dataset_base_path = ''
dataset = 'Inspec'
output_base_path = '../result/'
iteration = ''
gpu = 1
lr = 0.05
anneal_factor = 0.5
patience = 4
batch_size = 4
num_epochs = 100
threads = 12
param_selection_mode = False
use_tensorboard = False
no_dev = False
use_crf = True
rnn_layers = 3
hidden_size = 128
dropout = 0.3
word_dropout = 0.05
locked_dropout = 0.5
downsample_train = 0.0
not_in_memory = False

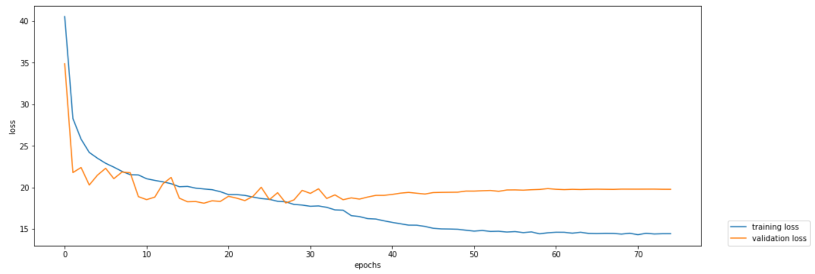
6500 s

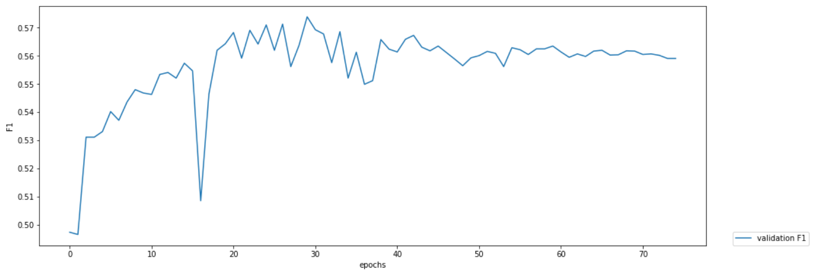
EPOCH 75 done: loss 14.4259 - lr 0.0001953
2021-06-24 00:29:32,972 DEV : loss 19.767724990844727 - score 0.5591

KEY tp: 3406 - fp: 2953 - fn: 1438 - precision: 0.5356 - recall: 0.7031 - f1-score: 0.6081

KEY tp: 3073 - fp: 3070 - fn: 1493 - precision: 0.5002 - recall: 0.6730 - f1-score: 0.5739
tensor(19.6010, device='cuda:0')

KEY tp: 7691 - fp: 5382 - fn: 2087 - precision: 0.5883 - recall: 0.7866 - f1-score: 0.6731
tensor(15.7998, device='cuda:0')



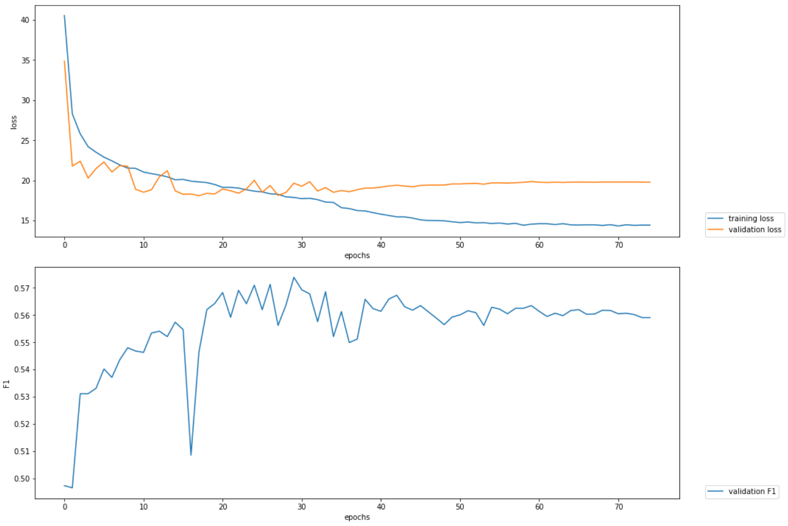


6400 s
EPOCH 72 done: loss 14.8274 - lr 0.0001953
2021-06-24 00:26:17,118 DEV : loss 19.666908264160156 - score 0.5605

KEY tp: 3334 - fp: 2870 - fn: 1510 - precision: 0.5374 - recall: 0.6883 - f1-score: 0.6035

KEY tp: 3040 - fp: 2971 - fn: 1526 - precision: 0.5057 - recall: 0.6658 - f1-score: 0.5748
tensor(19.5983, device='cuda:0')

KEY tp: 7467 - fp: 5357 - fn: 2311 - precision: 0.5823 - recall: 0.7637 - f1-score: 0.6607
tensor(16.4414, device='cuda:0')

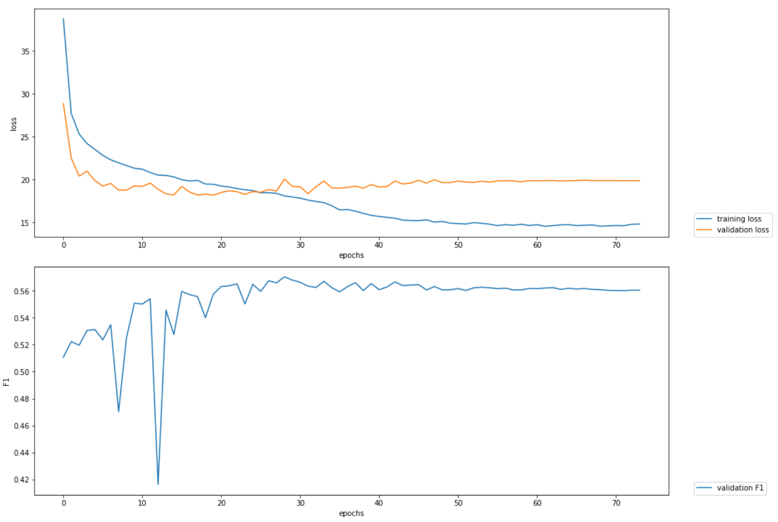


6500 s
EPOCH 74 done: loss 14.8263 - lr 0.0001953
2021-06-24 00:28:10,907 DEV : loss 19.899412155151367 - score 0.5605

KEY tp: 3434 - fp: 3128 - fn: 1410 - precision: 0.5233 - recall: 0.7089 - f1-score: 0.6021

KEY tp: 3128 - fp: 3274 - fn: 1438 - precision: 0.4886 - recall: 0.6851 - f1-score: 0.5704
tensor(20.0235, device='cuda:0')

KEY tp: 7691 - fp: 5803 - fn: 2087 - precision: 0.5700 - recall: 0.7866 - f1-score: 0.6610
tensor(16.5922, device='cuda:0')

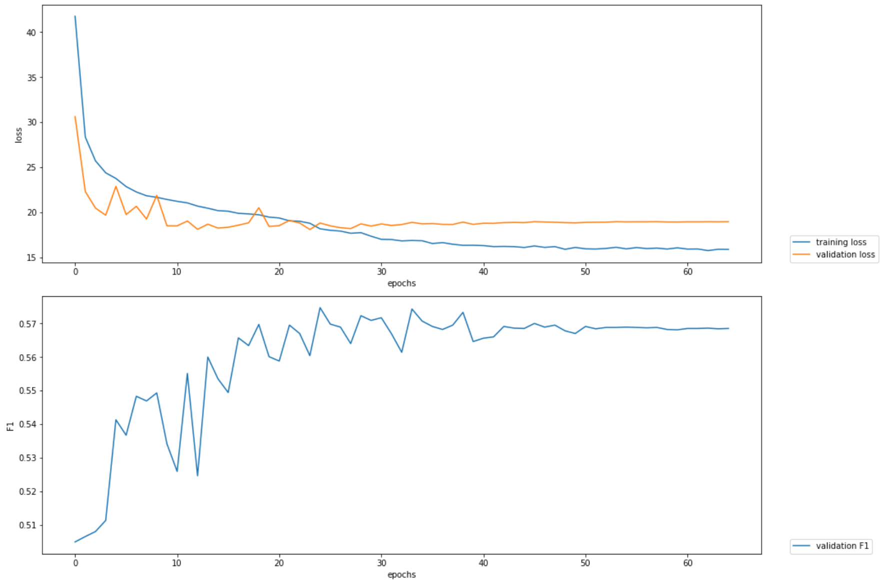


5800 s
EPOCH 65 done: loss 15.8822 - lr 0.0001953
2021-06-24 00:16:46,558 DEV : loss 18.955224990844727 - score 0.5685

KEY tp: 3200 - fp: 2645 - fn: 1644 - precision: 0.5475 - recall: 0.6606 - f1-score: 0.5987

KEY tp: 2959 - fp: 2772 - fn: 1607 - precision: 0.5163 - recall: 0.6481 - f1-score: 0.5747
tensor(18.7752, device='cuda:0')

KEY tp: 7295 - fp: 4869 - fn: 2483 - precision: 0.5997 - recall: 0.7461 - f1-score: 0.6649
tensor(16.0505, device='cuda:0')

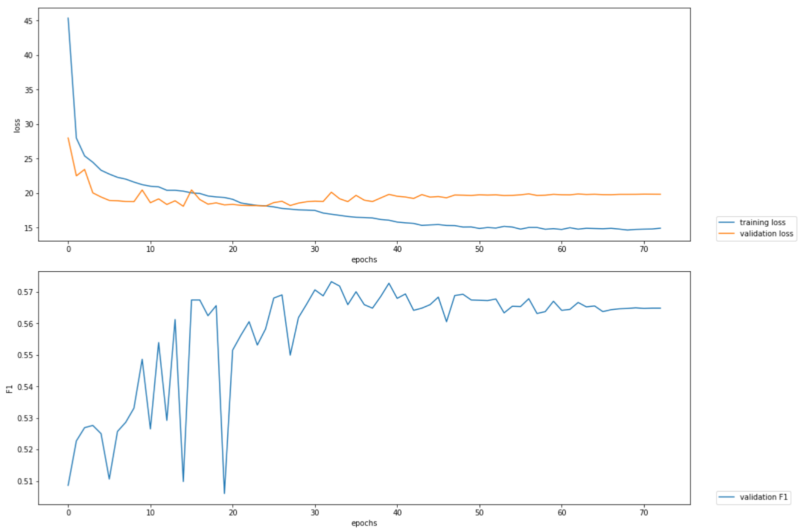


6300 s
EPOCH 73 done: loss 14.9303 - lr 0.0001953
2021-06-24 00:24:58,016 DEV : loss 19.840255737304688 - score 0.5648

KEY tp: 3386 - fp: 2960 - fn: 1458 - precision: 0.5336 - recall: 0.6990 - f1-score: 0.6052

KEY tp: 3077 - fp: 3093 - fn: 1489 - precision: 0.4987 - recall: 0.6739 - f1-score: 0.5732
tensor(20.0983, device='cuda:0')

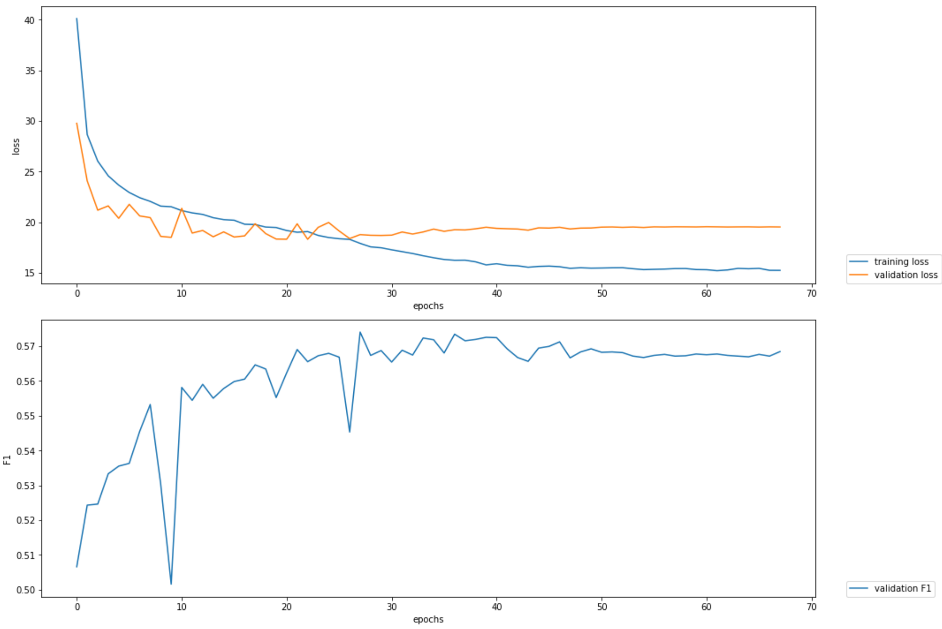
KEY tp: 7897 - fp: 5166 - fn: 1881 - precision: 0.6045 - recall: 0.8076 - f1-score: 0.6915
tensor(15.0742, device='cuda:0')



6000 s
EPOCH 68 done: loss 15.2412 - lr 0.0001953
2021-06-24 10:20:39,073 DEV : loss 19.528337478637695 - score 0.5684

KEY tp: 3105 - fp: 2410 - fn: 1739 - precision: 0.5630 - recall: 0.6410 - f1-score: 0.5995
KEY tp: 2847 - fp: 2506 - fn: 1719 - precision: 0.5319 - recall: 0.6235 - f1-score: 0.5740
tensor(18.7390, device='cuda:0')

KEY tp: 7250 - fp: 4148 - fn: 2528 - precision: 0.6361 - recall: 0.7415 - f1-score: 0.6847
tensor(14.9261, device='cuda:0')



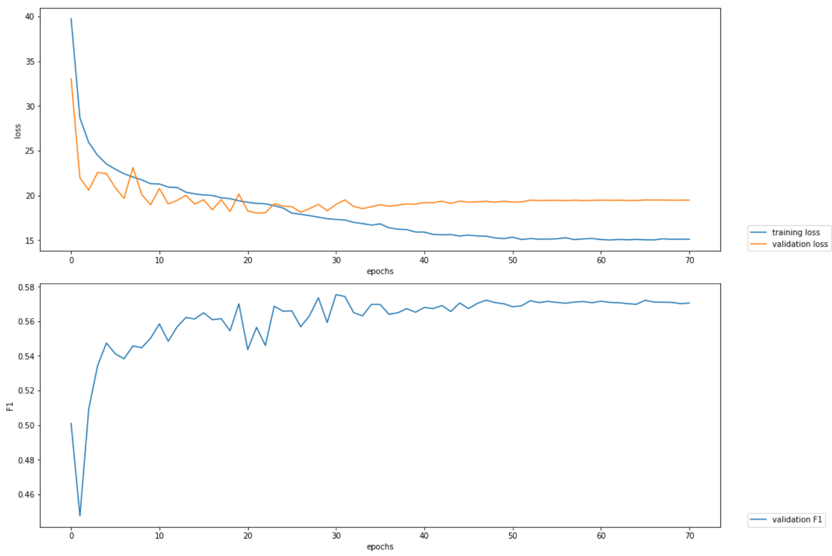
6250 s

EPOCH 71 done: loss 15.1045 - lr 0.0001953
2021-06-24 16:06:52,041 DEV : loss 19.47893524169922 - score 0.5705

KEY tp: 3189 - fp: 2554 - fn: 1655 - precision: 0.5553 - recall: 0.6583 - f1-score: 0.6024

KEY tp: 2912 - fp: 2644 - fn: 1654 - precision: 0.5241 - recall: 0.6378 - f1-score: 0.5754
tensor(18.9625, device='cuda:0')

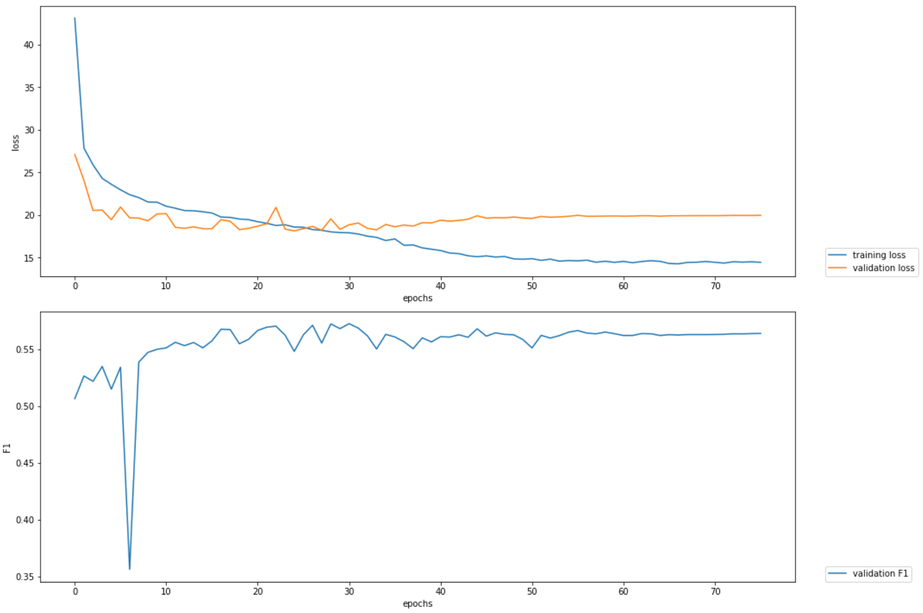
KEY tp: 7545 - fp: 4178 - fn: 2233 - precision: 0.6436 - recall: 0.7716 - f1-score: 0.7018
tensor(14.3419, device='cuda:0')



6500 s
EPOCH 76 done: loss 14.4617 - lr 0.0001953
2021-06-25 09:13:28,396 DEV : loss 19.97829246520996 - score 0.5641
KEY	tp: 3175 - fp: 2585 - fn: 1669 - precision: 0.5512 - recall: 0.6555 - f1-score: 0.5988

KEY	tp: 2906 - fp: 2677 - fn: 1660 - precision: 0.5205 - recall: 0.6364 - f1-score: 0.5727
tensor(18.8409, device='cuda:0')

KEY	tp: 7393 - fp: 4480 - fn: 2385 - precision: 0.6227 - recall: 0.7561 - f1-score: 0.6829
tensor(15.0114, device='cuda:0')

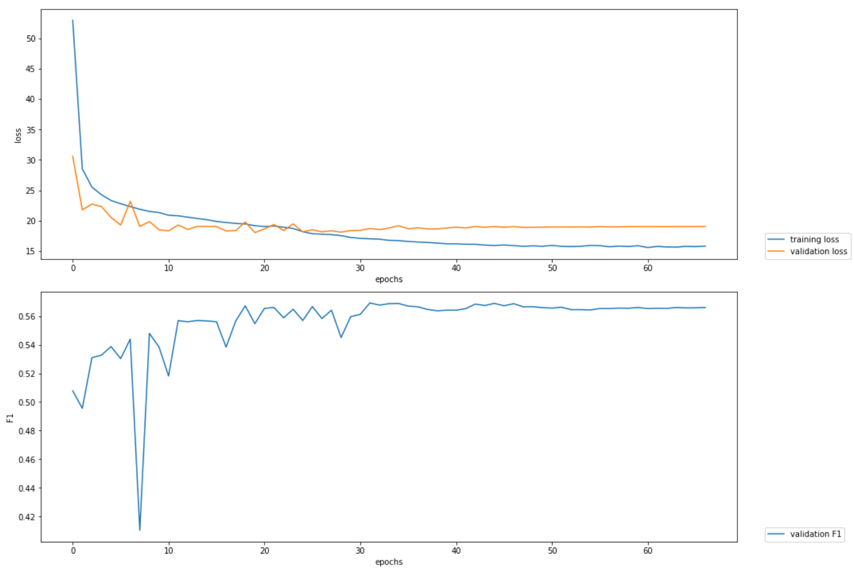


5800 s
EPOCH 67 done: loss 15.8382 - lr 0.0001953
2021-06-25 09:02:29,270 DEV : loss 19.059219360351562 - score 0.5661

KEY	tp: 3027 - fp: 2311 - fn: 1817 - precision: 0.5671 - recall: 0.6249 - f1-score: 0.5946

KEY	tp: 2763 - fp: 2378 - fn: 1803 - precision: 0.5374 - recall: 0.6051 - f1-score: 0.5693
tensor(18.6942, device='cuda:0')

KEY	tp: 7259 - fp: 3679 - fn: 2519 - precision: 0.6636 - recall: 0.7424 - f1-score: 0.7008
tensor(13.9809, device='cuda:0')

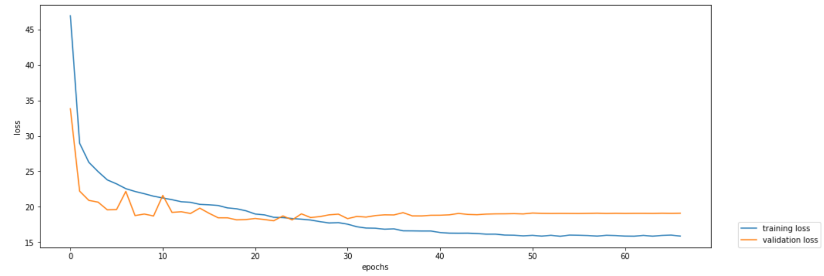


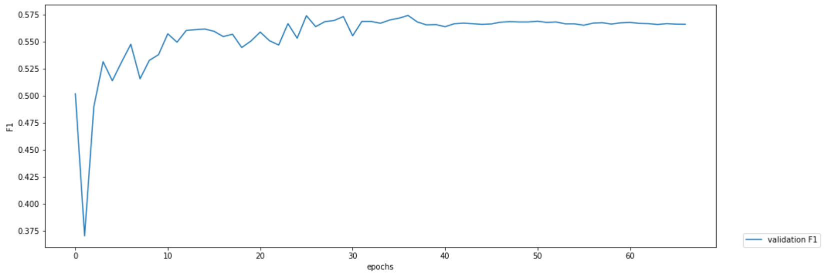
5800 s
EPOCH 67 done: loss 15.8840 - lr 0.0001953
2021-06-25 09:03:27,626 DEV : loss 19.10519790649414 - score 0.566

KEY tp: 3131 - fp: 2496 - fn: 1713 - precision: 0.5564 - recall: 0.6464 - f1-score: 0.5980

KEY tp: 2885 - fp: 2600 - fn: 1681 - precision: 0.5260 - recall: 0.6318 - f1-score: 0.5741
tensor(19.1374, device='cuda:0')

KEY tp: 7626 - fp: 3965 - fn: 2152 - precision: 0.6579 - recall: 0.7799 - f1-score: 0.7137
tensor(13.6252, device='cuda:0')



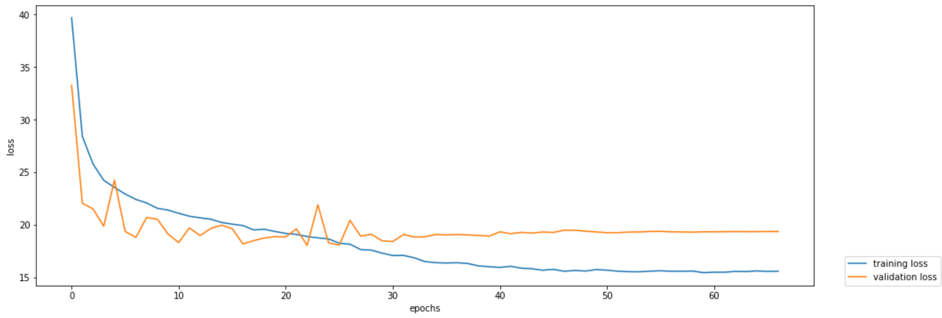


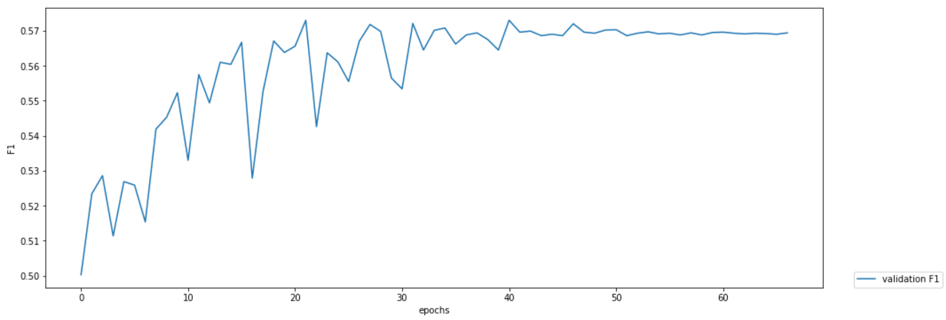
5900 s
EPOCH 67 done: loss 15.5933 - lr 0.0001953
2021-06-25 09:05:21,433 DEV : loss 19.364627838134766 - score 0.5694

KEY tp: 3367 - fp: 2969 - fn: 1477 - precision: 0.5314 - recall: 0.6951 - f1-score: 0.6023

KEY tp: 3089 - fp: 3127 - fn: 1477 - precision: 0.4969 - recall: 0.6765 - f1-score: 0.5730
tensor(19.5862, device='cuda:0')

KEY tp: 7306 - fp: 5875 - fn: 2472 - precision: 0.5543 - recall: 0.7472 - f1-score: 0.6364
tensor(17.7880, device='cuda:0')



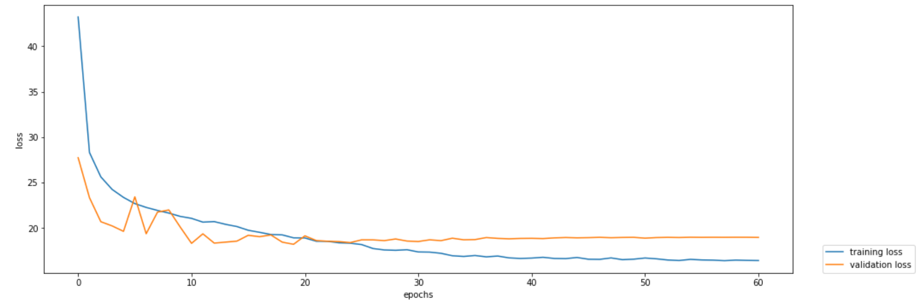


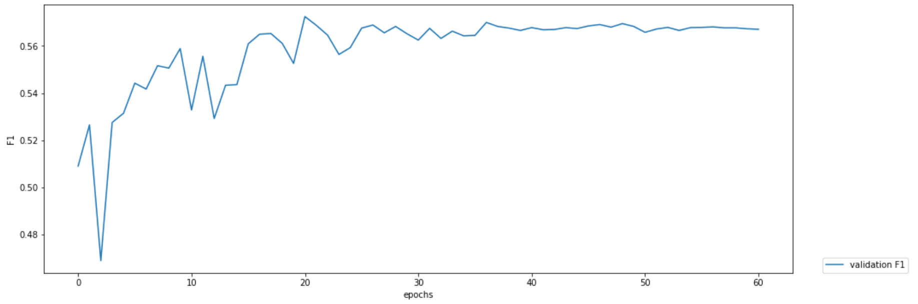
5500 s
EPOCH 61 done: loss 16.4111 - lr 0.0001953
2021-06-25 08:57:47,759 DEV : loss 18.957151412963867 - score 0.5671

KEY tp: 3247 - fp: 2735 - fn: 1597 - precision: 0.5428 - recall: 0.6703 - f1-score: 0.5999

KEY tp: 2971 - fp: 2842 - fn: 1595 - precision: 0.5111 - recall: 0.6507 - f1-score: 0.5725
tensor(19.0906, device='cuda:0')

KEY tp: 7175 - fp: 5258 - fn: 2603 - precision: 0.5771 - recall: 0.7338 - f1-score: 0.6461
tensor(17.1237, device='cuda:0')



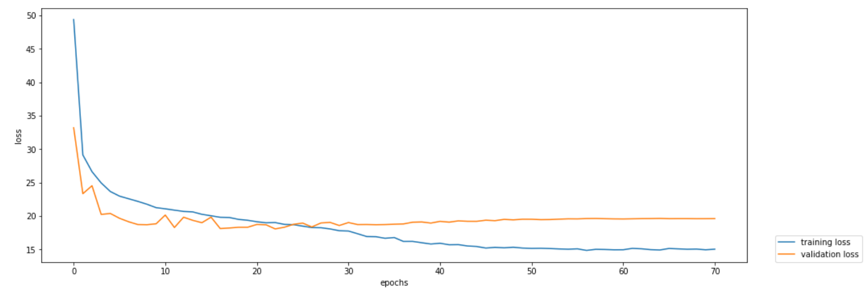


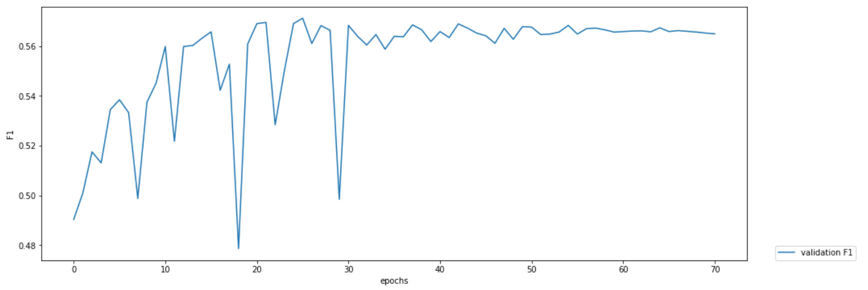
6150 s
EPOCH 71 done: loss 15.0421 - lr 0.0001953
2021-06-25 12:21:14,115 DEV : loss 19.607202529907227 - score 0.565

KEY tp: 3189 - fp: 2571 - fn: 1655 - precision: 0.5536 - recall: 0.6583 - f1-score: 0.6015

KEY tp: 2906 - fp: 2701 - fn: 1660 - precision: 0.5183 - recall: 0.6364 - f1-score: 0.5713
tensor(18.9203, device='cuda:0')

KEY tp: 7212 - fp: 4781 - fn: 2566 - precision: 0.6014 - recall: 0.7376 - f1-score: 0.6625
tensor(15.9900, device='cuda:0')



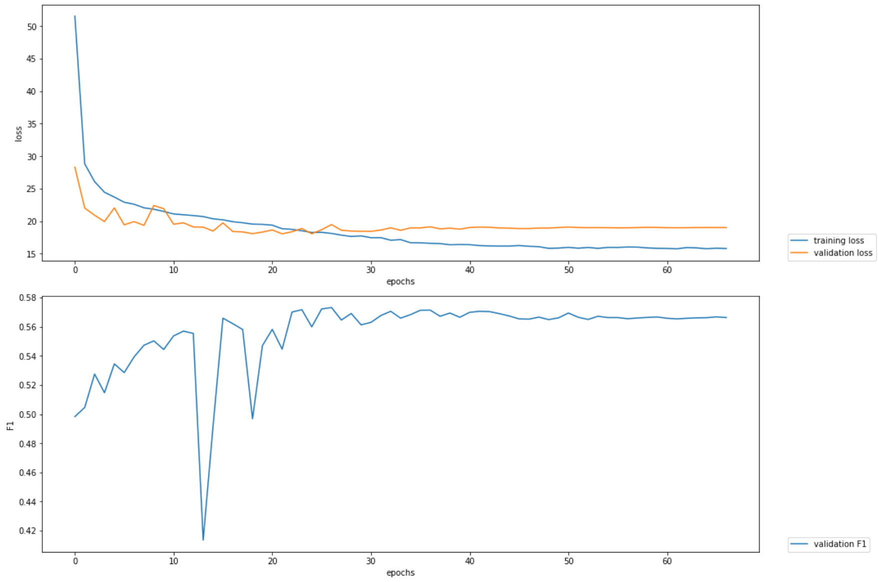


5900 s
EPOCH 67 done: loss 15.7771 - lr 0.0001953
2021-06-25 12:17:13,687 DEV : loss 19.006650924682617 - score 0.5663

KEY tp: 3397 - fp: 3029 - fn: 1447 - precision: 0.5286 - recall: 0.7013 - f1-score: 0.6028

KEY tp: 3095 - fp: 3138 - fn: 1471 - precision: 0.4966 - recall: 0.6778 - f1-score: 0.5732
tensor(19.4341, device='cuda:0')

KEY tp: 7664 - fp: 5592 - fn: 2114 - precision: 0.5782 - recall: 0.7838 - f1-score: 0.6655
tensor(16.5215, device='cuda:0')

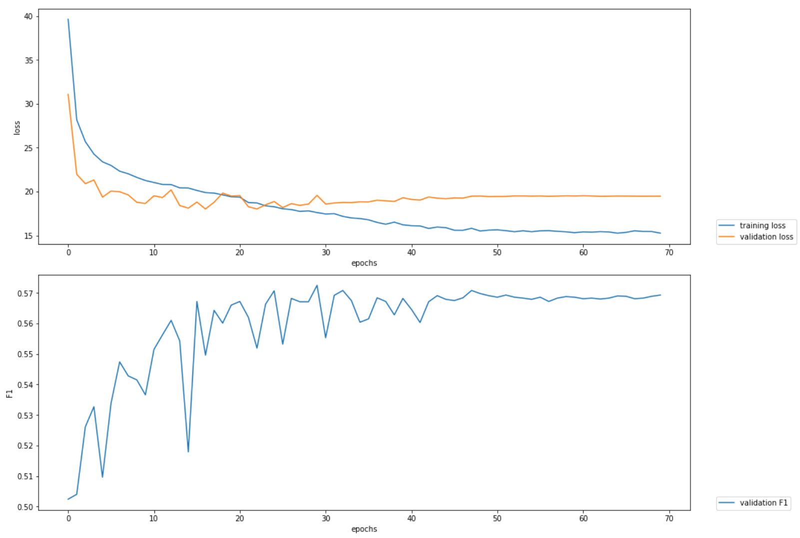


6200 s
EPOCH 70 done: loss 15.2640 - lr 0.0001953
2021-06-25 12:22:37,827 DEV : loss 19.471874237060547 - score 0.5693

KEY tp: 3382 - fp: 2963 - fn: 1462 - precision: 0.5330 - recall: 0.6982 - f1-score: 0.6045

KEY tp: 3053 - fp: 3046 - fn: 1513 - precision: 0.5006 - recall: 0.6686 - f1-score: 0.5725
tensor(19.5305, device='cuda:0')

KEY tp: 7728 - fp: 5211 - fn: 2050 - precision: 0.5973 - recall: 0.7903 - f1-score: 0.6804
tensor(15.6264, device='cuda:0')



Otros datasets

KDD 3x128 DO 0.3

embedding = 'Bert'
 embedding_path = ''
 dataset_base_path = ''
 dataset = 'Inspec'
 output_base_path = '../result/'
 iteration = ''
 gpu = 1
 lr = 0.05
 anneal_factor = 0.5
 patience = 4
 batch_size = 8
 num_epochs = 100
 threads = 12
 param_selection_mode = False
 use_tensorboard = False
 no_dev = False
 use_crf = True
 rnn_layers = 3
 hidden_size = 128
 dropout = 0.3
 word_dropout = 0.05
 locked_dropout = 0.5
 downsample_train = 0.0
 not_in_memory = False

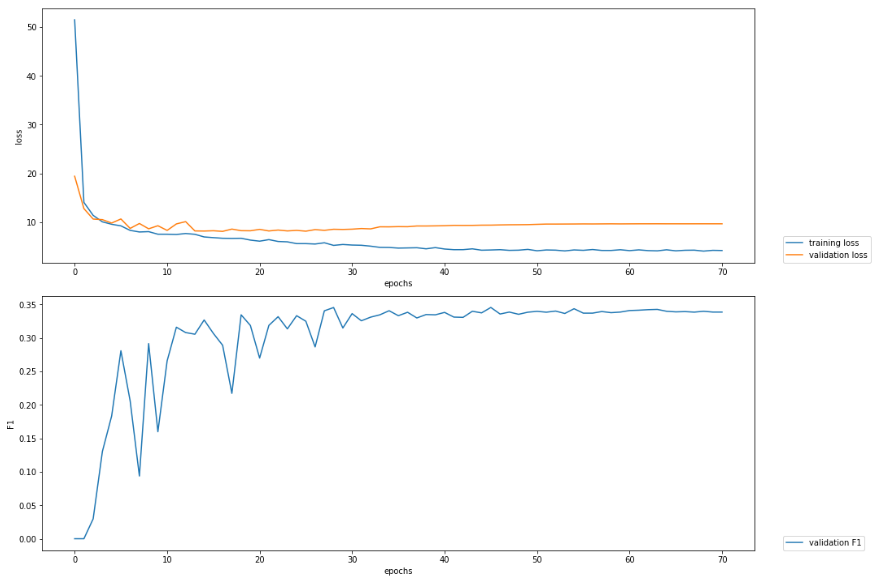
1300 s

EPOCH 71 done: loss 4.2330 - lr 0.0001953
2021-06-21 16:17:51,829 DEV : loss 9.699817657470703 - score 0.3387

KEY tp: 98 - fp: 175 - fn: 205 - precision: 0.3590 - recall: 0.3234 - f1-score: 0.3403

KEY tp: 141 - fp: 238 - fn: 296 - precision: 0.3720 - recall: 0.3227 - f1-score: 0.3456
tensor(9.5086, device='cuda:0')

KEY tp: 566 - fp: 204 - fn: 237 - precision: 0.7351 - recall: 0.7049 - f1-score: 0.7196
tensor(2.7916, device='cuda:0')



KDD 3x64

embedding = 'Bert'
embedding_path = ''
dataset_base_path = ''
dataset = 'Inspec'
output_base_path = '../result/'
iteration = ''
gpu = 1
lr = 0.05
anneal_factor = 0.5
patience = 4
batch_size = 8
num_epochs = 100
threads = 12
param_selection_mode = False
use_tensorboard = False
no_dev = False
use_crf = True
rnn_layers = 3
hidden_size = 64
dropout = 0.3
word_dropout = 0.05
locked_dropout = 0.5
downsample_train = 0.0
not_in_memory = False

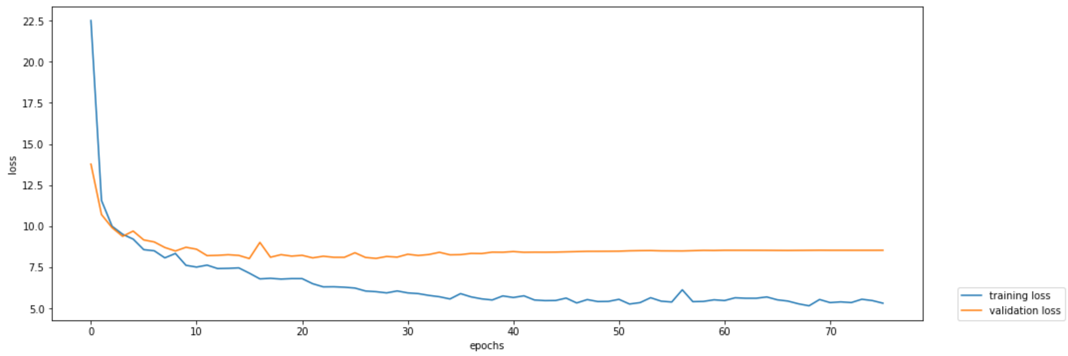
EPOCH 76 done: loss 5.3165 - lr 0.0001953
2021-06-21 16:48:28,335 DEV : loss 8.536468505859375 - score 0.3226

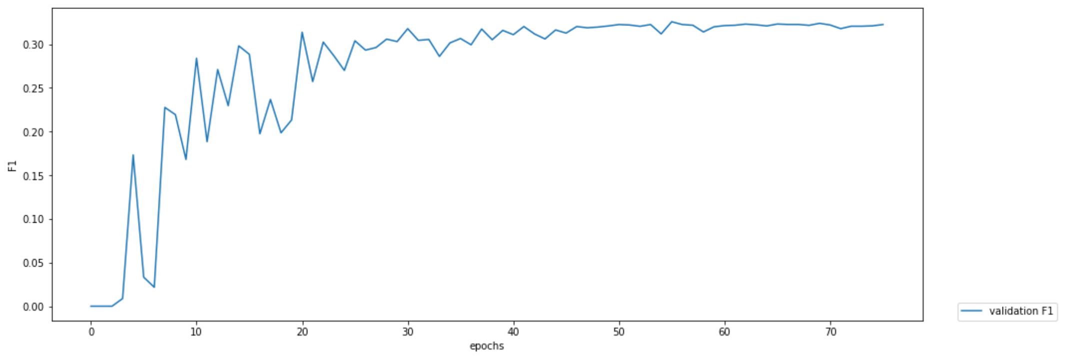
1500 s

KEY tp: 60 - fp: 111 - fn: 243 - precision: 0.3509 - recall: 0.1980 - f1-score: 0.2532

KEY tp: 116 - fp: 159 - fn: 321 - precision: 0.4218 - recall: 0.2654 - f1-score: 0.3258
tensor(8.5588, device='cuda:0')

KEY tp: 339 - fp: 201 - fn: 464 - precision: 0.6278 - recall: 0.4222 - f1-score: 0.5048
tensor(3.9160, device='cuda:0')





INSPEC MIO

embedding = 'Bert'
embedding_path = ''
dataset_base_path = ''
dataset = 'Inspec'
output_base_path = '../result/'
iteration = ''
gpu = 1
lr = 0.1
anneal_factor = 0.5
patience = 4
batch_size = 16
num_epochs = 100
threads = 12
param_selection_mode = False
use_tensorboard = False
no_dev = False
use_crf = True
rnn_layers = 3
hidden_size = 128
dropout = 0.3
word_dropout = 0.05
locked_dropout = 0.5
downsample_train = 0.0
not_in_memory = False

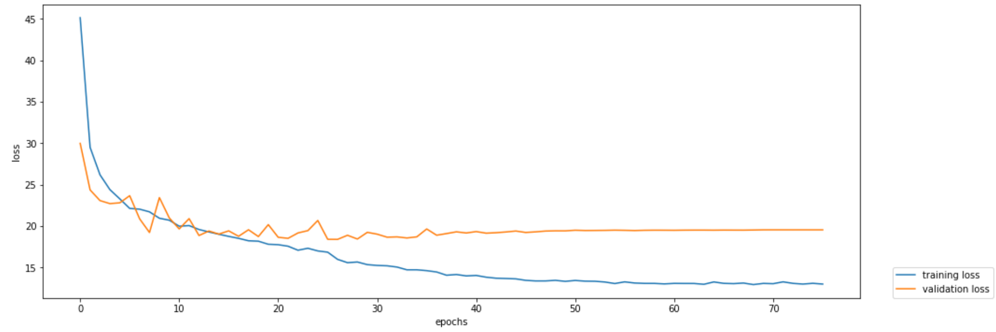
5600 s

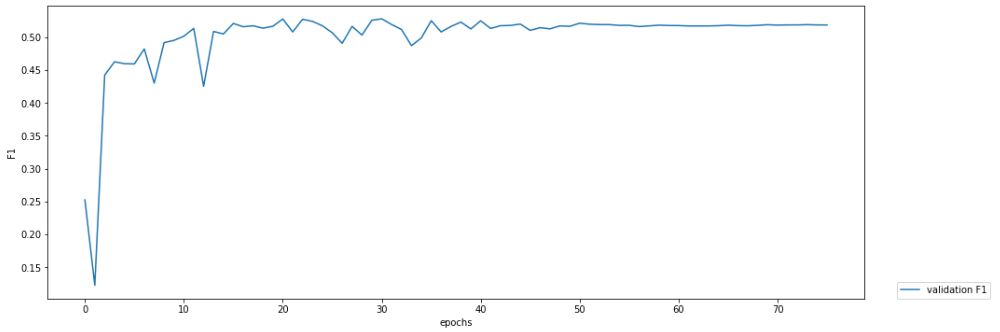
EPOCH 76 done: loss 12.9887 - lr 0.0001953
2021-06-22 10:12:40,254 DEV : loss 19.53164291381836 - score 0.5186

KEY tp: 2570 - fp: 2321 - fn: 1670 - precision: 0.5255 - recall: 0.6061 - f1-score: 0.5629

KEY tp: 2357 - fp: 2469 - fn: 1745 - precision: 0.4884 - recall: 0.5746 - f1-score: 0.5280
tensor(19.2053, device='cuda:0')

KEY tp: 6306 - fp: 3882 - fn: 2014 - precision: 0.6190 - recall: 0.7579 - f1-score: 0.6814
tensor(12.7253, device='cuda:0')





KP20K 4000 files

embedding = 'Bert'
embedding_path = ''
dataset_base_path = ''
dataset = 'Inspec'
output_base_path = '../result/'
iteration = ''
gpu = 1
lr = 0.05
anneal_factor = 0.5
patience = 4
batch_size = 4
num_epochs = 100
threads = 12
param_selection_mode = False
use_tensorboard = False
no_dev = False
use_crf = True
rnn_layers = 3
hidden_size = 128
dropout = 0.3
word_dropout = 0.05
locked_dropout = 0.5
downsample_train = 0.0
not_in_memory = False

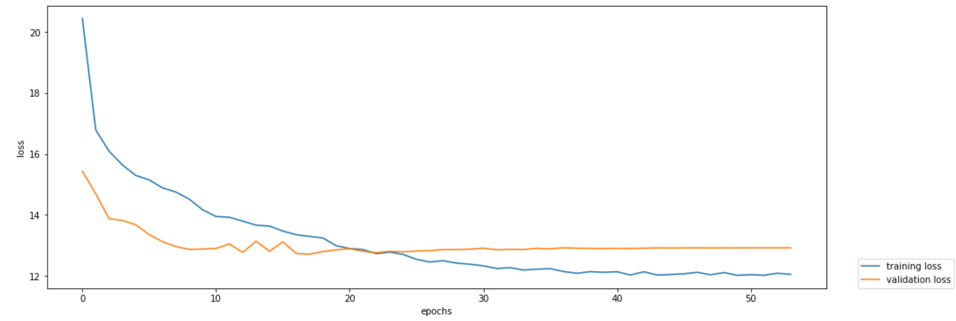
6h30min

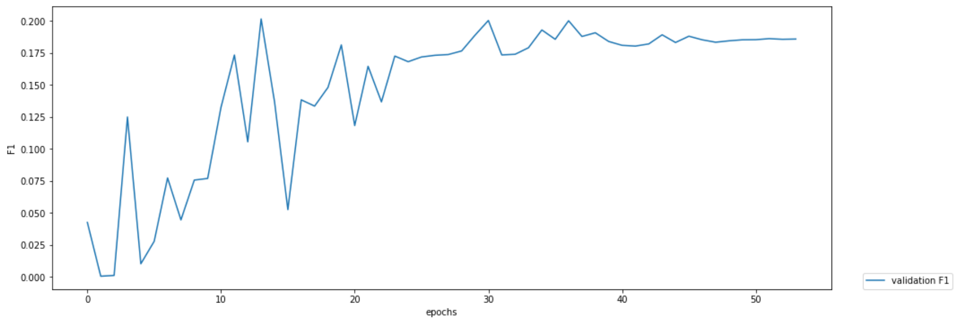
EPOCH 54 done: loss 12.0506 - lr 0.0001953
2021-06-22 17:50:37,893 DEV : loss 12.9235200881958 - score 0.1858

KEY tp: 499 - fp: 1008 - fn: 3391 - precision: 0.3311 - recall: 0.1283 - f1-score: 0.1849

KEY tp: 522 - fp: 969 - fn: 3168 - precision: 0.3501 - recall: 0.1415 - f1-score: 0.2015
tensor(13.2679, device='cuda:0')

EY tp: 1631 - fp: 1745 - fn: 6237 - precision: 0.4831 - recall: 0.2073 - f1-score: 0.2901
tensor(12.4791, device='cuda:0')





INSPEC MM 500

3900 s
EPOCH 63 done: loss 16.3039 - lr 0.0001953
2021-06-29 16:40:45,397 DEV : loss 20.052637100219727 - score 0.5579

KEY tp: 3238 - fp: 2804 - fn: 1606 - precision: 0.5359 - recall: 0.6685 - f1-score: 0.5949

KEY tp: 2915 - fp: 2926 - fn: 1651 - precision: 0.4991 - recall: 0.6384 - f1-score: 0.5602
tensor(20.2911, device='cuda:0')

5000 s
EPOCH 82 done: loss 14.0935 - lr 0.0001953
2021-07-06 15:36:02,531 DEV : loss 20.4657039642334 - score 0.5523

KEY tp: 3200 - fp: 2788 - fn: 1644 - precision: 0.5344 - recall: 0.6606 - f1-score: 0.5908

KEY tp: 2880 - fp: 2800 - fn: 1686 - precision: 0.5070 - recall: 0.6307 - f1-score: 0.5622
tensor(20.1844, device='cuda:0')

3500 s
EPOCH 70 done: loss 15.5852 - lr 0.0001953
2021-07-06 15:12:50,016 DEV : loss 19.765214920043945 - score 0.5438

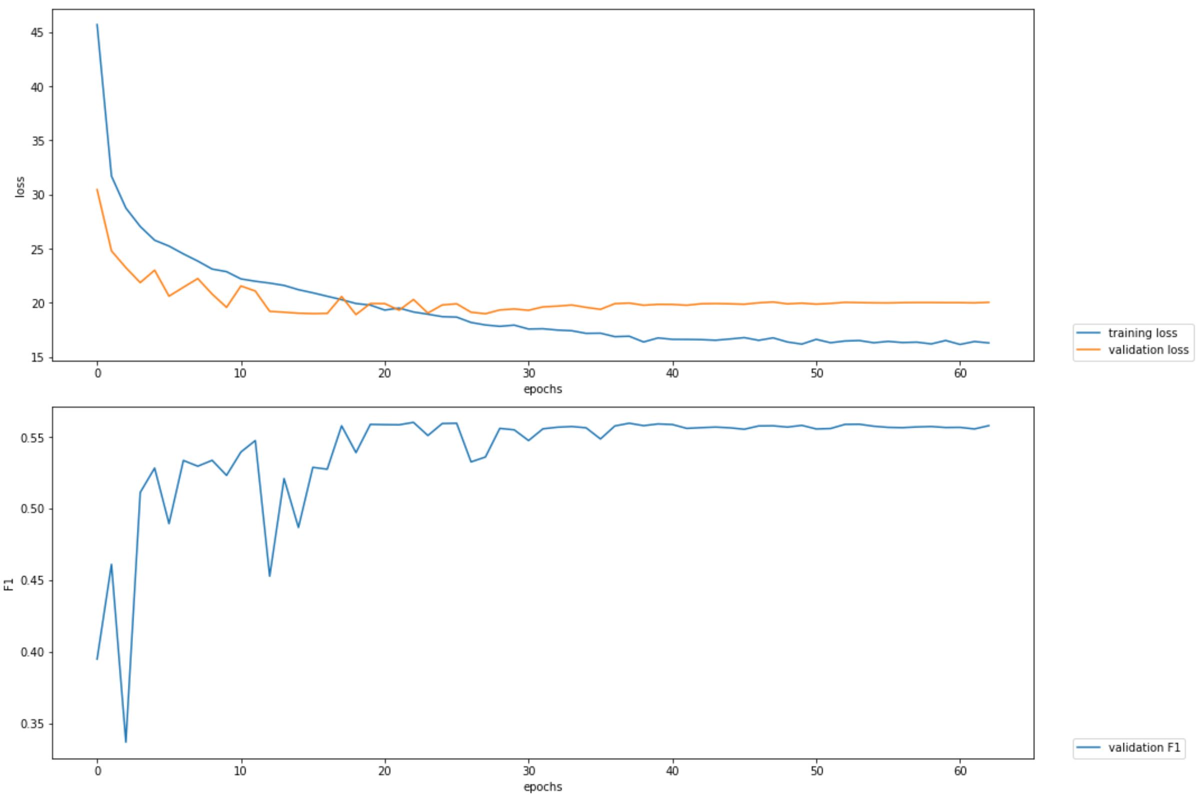
KEY tp: 3134 - fp: 2562 - fn: 1710 - precision: 0.5502 - recall: 0.6470 - f1-score: 0.5947

KEY tp: 2825 - fp: 2734 - fn: 1741 - precision: 0.5082 - recall: 0.6187 - f1-score: 0.5580
tensor(19.3977, device='cuda:0')

4700 s
EPOCH 69 done: loss 15.5729 - lr 0.0001953
2021-07-06 16:23:08,212 DEV : loss 20.648439407348633 - score 0.5563

KEY tp: 3106 - fp: 2592 - fn: 1738 - precision: 0.5451 - recall: 0.6412 - f1-score: 0.5893

KEY tp: 2814 - fp: 2645 - fn: 1752 - precision: 0.5155 - recall: 0.6163 - f1-score: 0.5614
tensor(20.2689, device='cuda:0')



INSPEC MM 250

2250 s
EPOCH 66 done: loss 16.2712 - lr 0.0001953
2021-06-29 15:23:10,697 DEV : loss 23.15253257751465 - score 0.539

KEY tp: 3190 - fp: 2968 - fn: 1654 - precision: 0.5180 - recall: 0.6585 - f1-score: 0.5799

KEY tp: 2924 - fp: 3170 - fn: 1642 - precision: 0.4798 - recall: 0.6404 - f1-score: 0.5486
tensor(22.9820, device='cuda:0')

KEY tp: 2012 - fp: 1136 - fn: 520 - precision: 0.6391 - recall: 0.7946 - f1-score: 0.7085
tensor(14.8378, device='cuda:0')

2600 s
EPOCH 78 done: loss 13.4851 - lr 0.0001953
2021-07-07 08:24:53,779 DEV : loss 22.9456729888916 - score 0.5366

KEY tp: 3112 - fp: 2830 - fn: 1732 - precision: 0.5237 - recall: 0.6424 - f1-score: 0.5770

KEY tp: 2862 - fp: 3048 - fn: 1704 - precision: 0.4843 - recall: 0.6268 - f1-score: 0.5464
tensor(22.7037, device='cuda:0')

2500 s
EPOCH 79 done: loss 12.8402 - lr 0.0001953
2021-07-07 08:23:52,858 DEV : loss 23.8172550201416 - score 0.5369

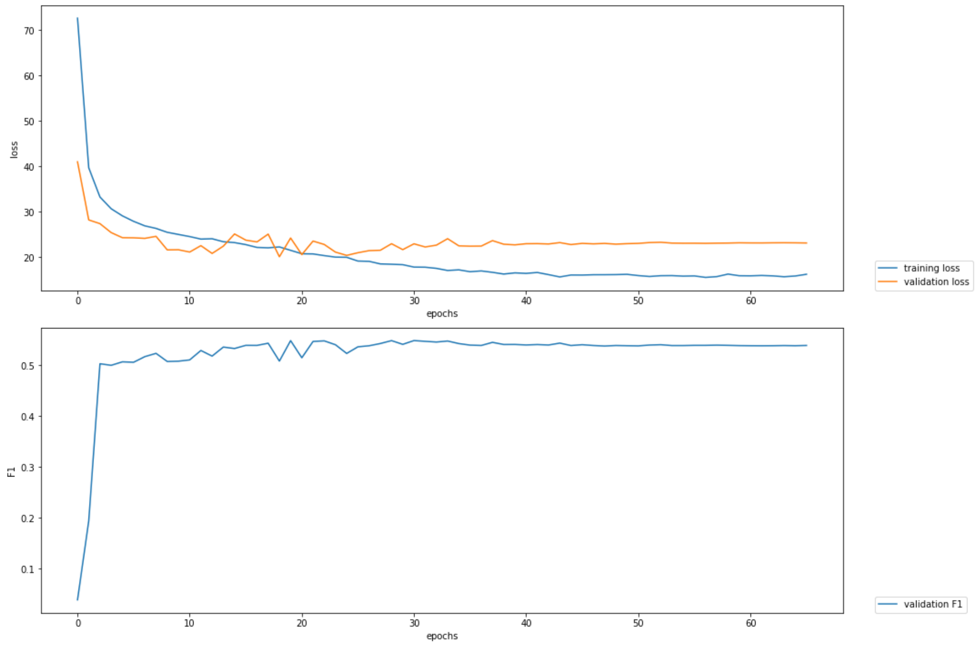
KEY tp: 3140 - fp: 2794 - fn: 1704 - precision: 0.5292 - recall: 0.6482 - f1-score: 0.5827

KEY tp: 2862 - fp: 2963 - fn: 1704 - precision: 0.4913 - recall: 0.6268 - f1-score: 0.5509
tensor(22.3766, device='cuda:0')

2550 s
EPOCH 79 done: loss 12.8402 - lr 0.0001953
2021-07-07 08:23:52,858 DEV : loss 23.8172550201416 - score 0.5369

KEY tp: 3140 - fp: 2794 - fn: 1704 - precision: 0.5292 - recall: 0.6482 - f1-score: 0.5827

KEY tp: 2862 - fp: 2963 - fn: 1704 - precision: 0.4913 - recall: 0.6268 - f1-score: 0.5509
tensor(22.3766, device='cuda:0')



INSPEC MM 100

1
2300 s
EPOCH 74 done: loss 13.1583 - lr 0.0001953
2021-07-07 12:33:44,167 DEV : loss 26.208282470703125 - score 0.5127

KEY tp: 3260 - fp: 3357 - fn: 1584 - precision: 0.4927 - recall: 0.6730 - f1-score: 0.5689

KEY tp: 2981 - fp: 3647 - fn: 1585 - precision: 0.4498 - recall: 0.6529 - f1-score: 0.5326
tensor(27.0394, device='cuda:0')

2
1600 s
EPOCH 73 done: loss 17.5300 - lr 0.0001953
2021-07-07 12:06:14,661 DEV : loss 26.83908462524414 - score 0.5216

KEY tp: 2972 - fp: 2944 - fn: 1872 - precision: 0.5024 - recall: 0.6135 - f1-score: 0.5524

KEY tp: 2750 - fp: 3132 - fn: 1816 - precision: 0.4675 - recall: 0.6023 - f1-score: 0.5264

tensor(26.9695, device='cuda:0')

3
2200 s
EPOCH 74 done: loss 12.5496 - lr 0.0001953
2021-07-07 12:33:01,571 DEV : loss 26.74735450744629 - score 0.5098

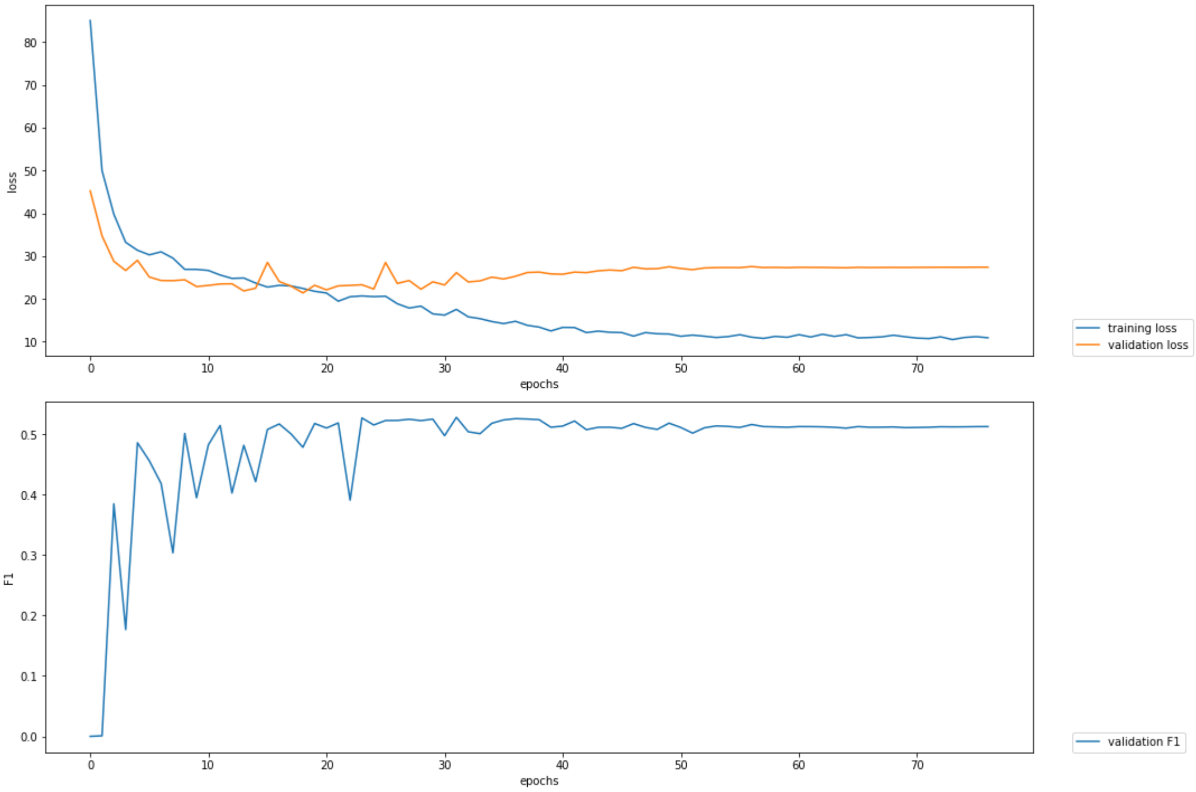
KEY tp: 3206 - fp: 3244 - fn: 1638 - precision: 0.4971 - recall: 0.6618 - f1-score: 0.5677

KEY tp: 2925 - fp: 3452 - fn: 1641 - precision: 0.4587 - recall: 0.6406 - f1-score: 0.5346
tensor(24.7690, device='cuda:0')

4
2400 s
EPOCH 77 done: loss 10.9175 - lr 0.0001953
2021-07-07 12:35:55,041 DEV : loss 27.409543991088867 - score 0.5127

KEY tp: 3125 - fp: 3370 - fn: 1719 - precision: 0.4811 - recall: 0.6451 - f1-score: 0.5512

KEY tp: 2885 - fp: 3476 - fn: 1681 - precision: 0.4535 - recall: 0.6318 - f1-score: 0.5280
tensor(26.1165, device='cuda:0')



INSPEC MM 50

1
1400 s
EPOCH 81 done: loss 11.0219 - lr 0.0001953
2021-07-06 08:35:58,192 DEV : loss 31.245080947875977 - score 0.4937

KEY tp: 3091 - fp: 3309 - fn: 1753 - precision: 0.4830 - recall: 0.6381 - f1-score: 0.5498

KEY tp: 2859 - fp: 3525 - fn: 1707 - precision: 0.4478 - recall: 0.6261 - f1-score: 0.5222
tensor(28.8116, device='cuda:0')

2
1200 s
EPOCH 69 done: loss 18.3052 - lr 0.0001953
2021-07-05 16:38:47,609 DEV : loss 29.669492721557617 - score 0.4868

KEY tp: 2998 - fp: 3295 - fn: 1846 - precision: 0.4764 - recall: 0.6189 - f1-score: 0.5384

KEY tp: 2710 - fp: 3558 - fn: 1856 - precision: 0.4324 - recall: 0.5935 - f1-score: 0.5003
tensor(31.1330, device='cuda:0')

3
1450 s
EPOCH 80 done: loss 11.4376 - lr 0.0001953
2021-07-05 16:42:52,478 DEV : loss 31.279273986816406 - score 0.4891

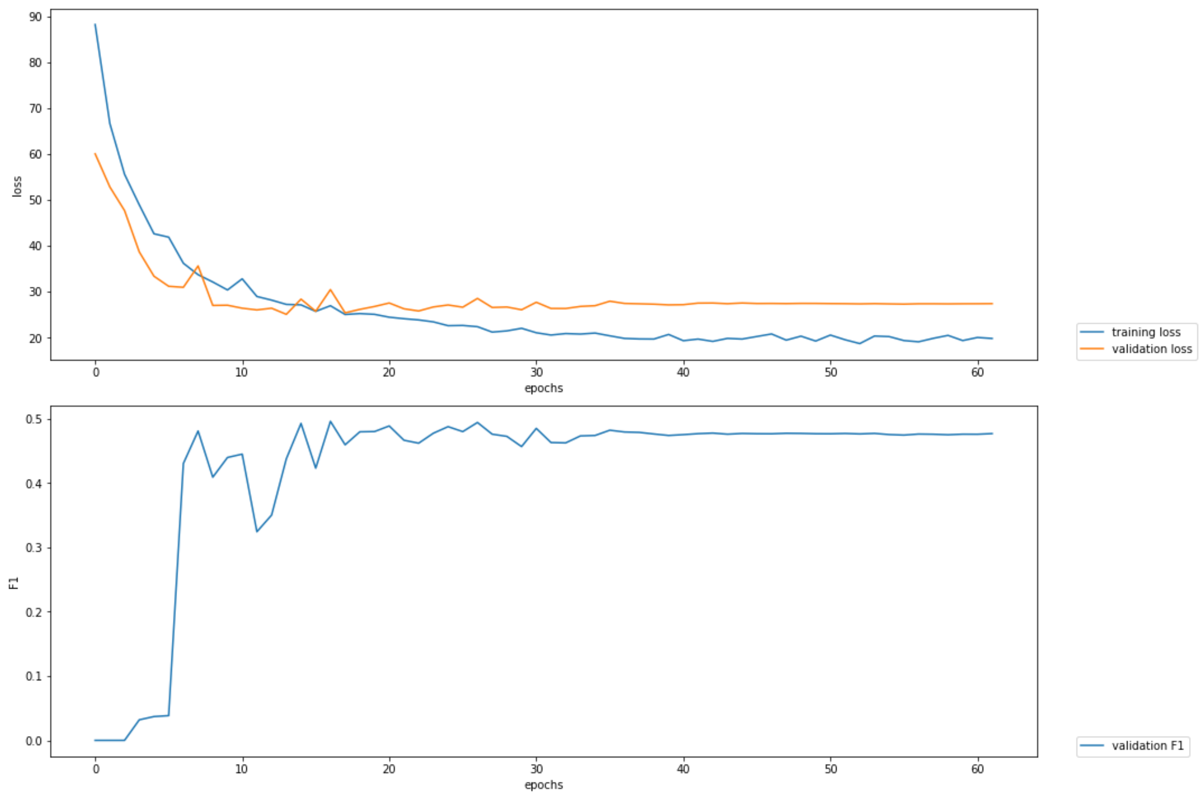
KEY tp: 3166 - fp: 3403 - fn: 1678 - precision: 0.4820 - recall: 0.6536 - f1-score: 0.5548

KEY tp: 2855 - fp: 3665 - fn: 1711 - precision: 0.4379 - recall: 0.6253 - f1-score: 0.5151
tensor(28.8084, device='cuda:0')

4
1350 s
EPOCH 78 done: loss 12.1858 - lr 0.0001953
2021-07-06 09:03:13,903 DEV : loss 30.21949005126953 - score 0.4899

KEY tp: 3216 - fp: 3857 - fn: 1628 - precision: 0.4547 - recall: 0.6639 - f1-score: 0.5397

KEY tp: 2950 - fp: 4011 - fn: 1616 - precision: 0.4238 - recall: 0.6461 - f1-score: 0.5118
tensor(31.3720, device='cuda:0')



INSPEC MM 25

1
1050 s
EPOCH 66 done: loss 19.7409 - lr 0.0001953
2021-07-06 08:00:30,456 DEV : loss 32.03760528564453 - score 0.4133

KEY tp: 2434 - fp: 3024 - fn: 2410 - precision: 0.4460 - recall: 0.5025 - f1-score: 0.4725

KEY tp: 2182 - fp: 3177 - fn: 2384 - precision: 0.4072 - recall: 0.4779 - f1-score: 0.4397
tensor(31.5628, device='cuda:0')

2
1200 s
EPOCH 78 done: loss 24.3124 - lr 0.0001953
2021-07-06 08:03:25,078 DEV : loss 33.69823455810547 - score 0.4102

KEY tp: 2591 - fp: 3112 - fn: 2253 - precision: 0.4543 - recall: 0.5349 - f1-score: 0.4913

KEY tp: 2324 - fp: 3270 - fn: 2242 - precision: 0.4154 - recall: 0.5090 - f1-score: 0.4575
tensor(33.3086, device='cuda:0')

3
2000 s
EPOCH 66 done: loss 20.7567 - lr 0.0001953
2021-07-06 08:17:30,844 DEV : loss 31.707260131835938 - score 0.4179

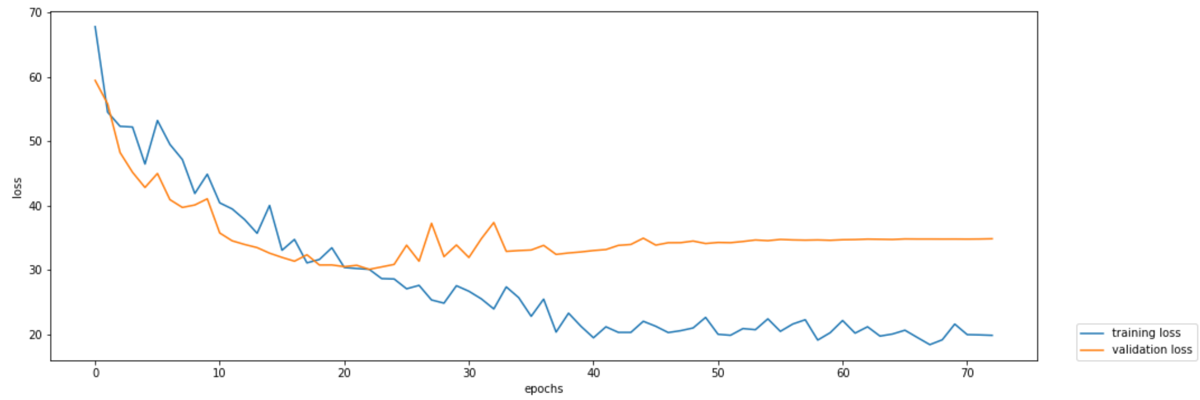
KEY tp: 2868 - fp: 3520 - fn: 1976 - precision: 0.4490 - recall: 0.5921 - f1-score: 0.5107

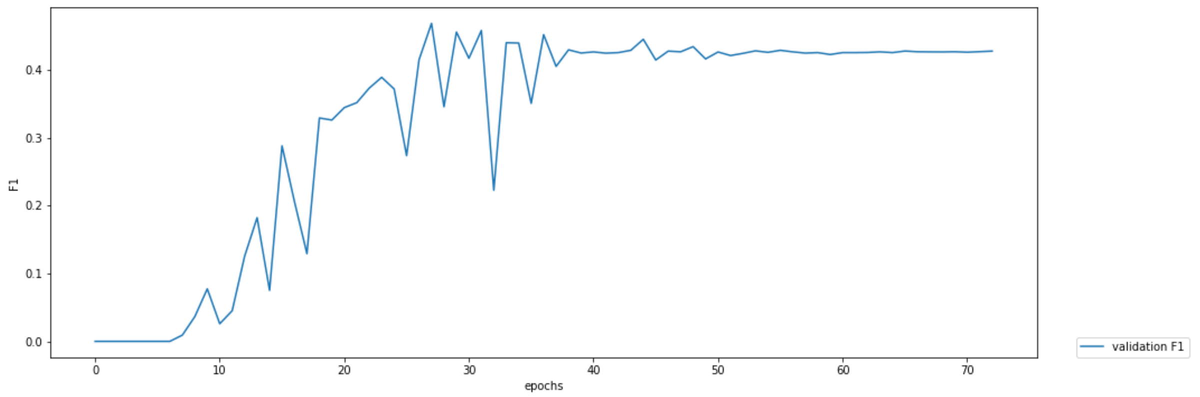
KEY tp: 2604 - fp: 3810 - fn: 1962 - precision: 0.4060 - recall: 0.5703 - f1-score: 0.4743
tensor(34.3665, device='cuda:0')

4
1150 s
EPOCH 73 done: loss 19.7957 - lr 0.0001953
2021-07-06 08:25:40,413 DEV : loss 34.81412124633789 - score 0.4274

KEY tp: 2777 - fp: 3358 - fn: 2067 - precision: 0.4526 - recall: 0.5733 - f1-score: 0.5059

KEY tp: 2493 - fp: 3588 - fn: 2073 - precision: 0.4100 - recall: 0.5460 - f1-score: 0.4683
tensor(37.1726, device='cuda:0')





KDD

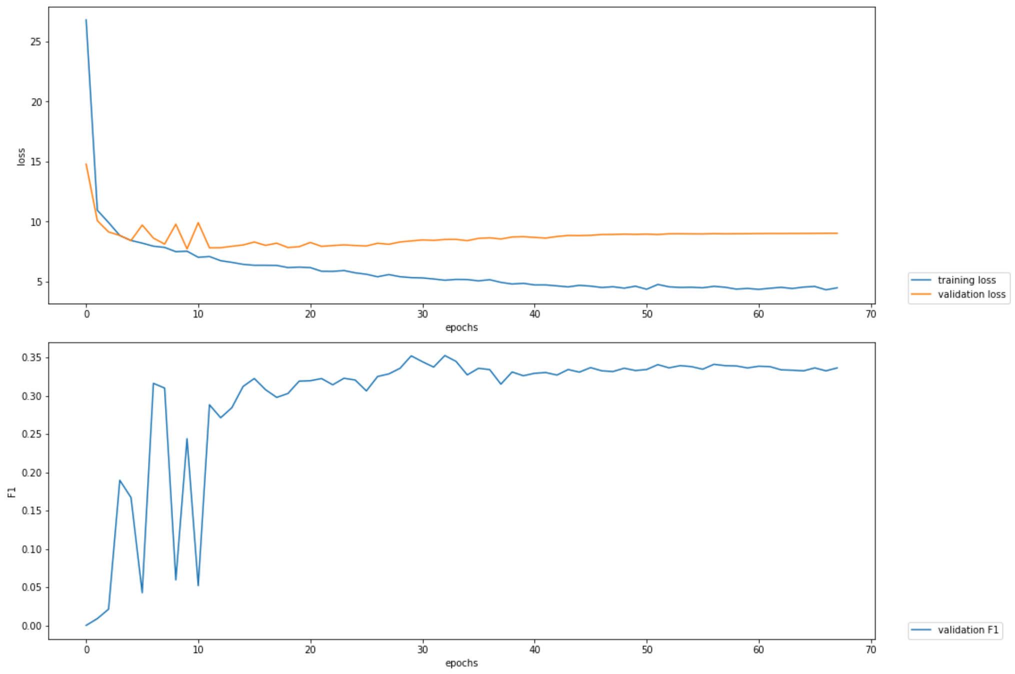
4000 s

EPOCH 68 done: loss 4.4875 - lr 0.0001953
2021-06-30 16:26:55,533 DEV : loss 9.023460388183594 - score 0.3363

KEY tp: 94 - fp: 160 - fn: 209 - precision: 0.3701 - recall: 0.3102 - f1-score: 0.3375

KEY tp: 144 - fp: 236 - fn: 293 - precision: 0.3789 - recall: 0.3295 - f1-score: 0.3525
tensor(8.6525, device='cuda:0')

KEY tp: 455 - fp: 286 - fn: 346 - precision: 0.6140 - recall: 0.5680 - f1-score: 0.5901
tensor(3.7957, device='cuda:0')

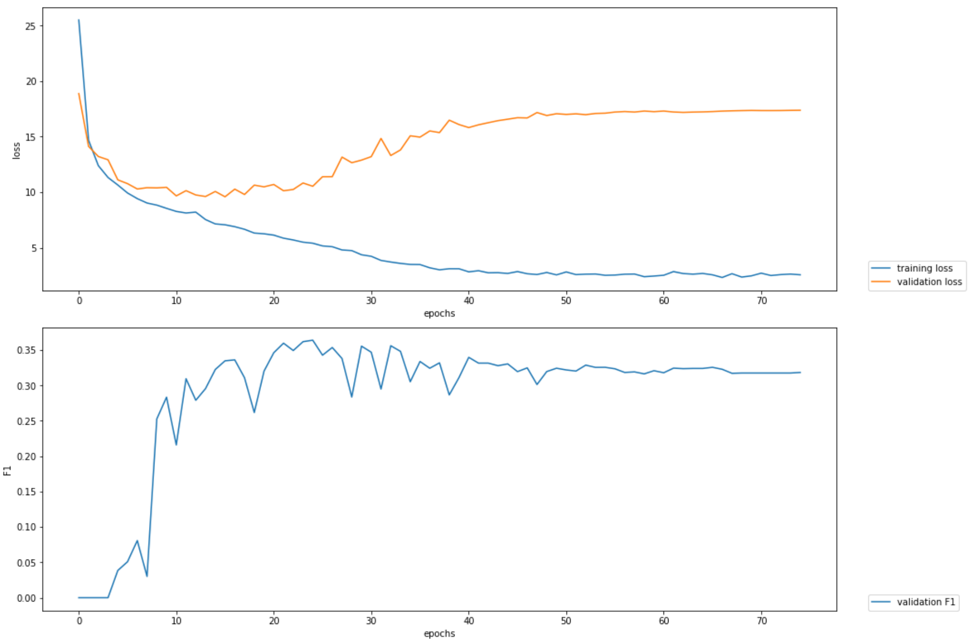


2200 s
EPOCH 75 done: loss 2.5680 - lr 0.0001953
2021-07-01 09:10:06,660 DEV : loss 17.38059425354004 - score 0.3182

KEY tp: 99 - fp: 176 - fn: 204 - precision: 0.3600 - recall: 0.3267 - f1-score: 0.3426

KEY tp: 157 - fp: 269 - fn: 280 - precision: 0.3685 - recall: 0.3593 - f1-score: 0.3638
tensor(10.7368, device='cuda:0')

KEY tp: 578 - fp: 237 - fn: 223 - precision: 0.7092 - recall: 0.7216 - f1-score: 0.7153
tensor(3.0485, device='cuda:0')



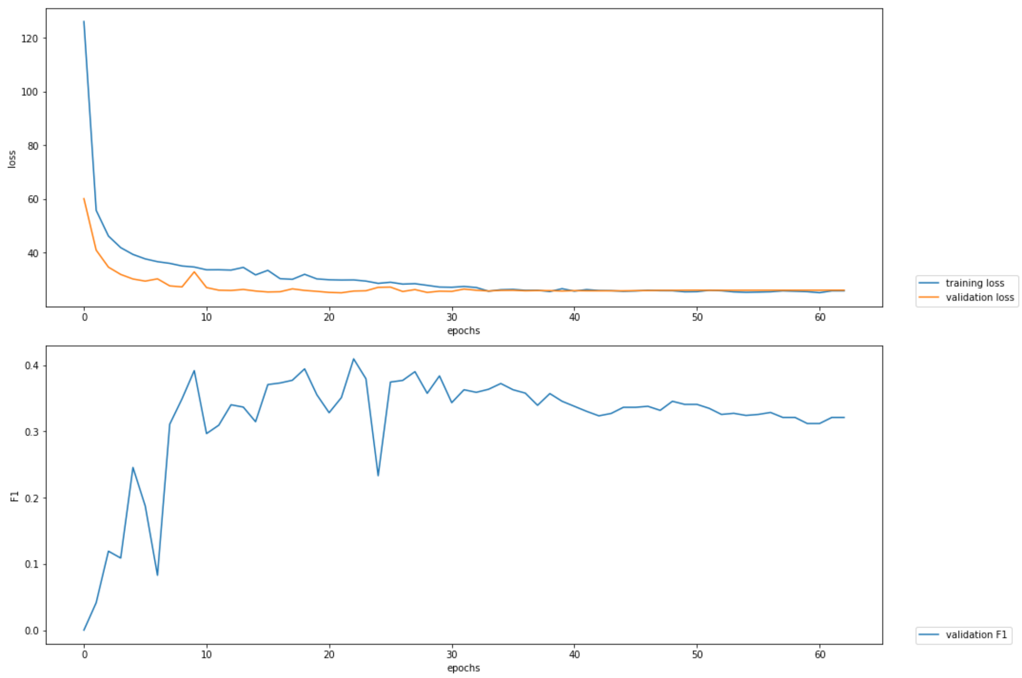
SE2010

800 s
EPOCH 63 done: loss 25.8596 - lr 0.0001953
2021-06-30 15:49:31,555 DEV : loss 26.108646392822266 - score 0.3211

KEY tp: 304 - fp: 590 - fn: 680 - precision: 0.3400 - recall: 0.3089 - f1-score: 0.3237

KEY tp: 43 - fp: 45 - fn: 79 - precision: 0.4886 - recall: 0.3525 - f1-score: 0.4095
tensor(25.8096, device='cuda:0')

KEY tp: 517 - fp: 573 - fn: 789 - precision: 0.4743 - recall: 0.3959 - f1-score: 0.4316
tensor(26.4052, device='cuda:0')



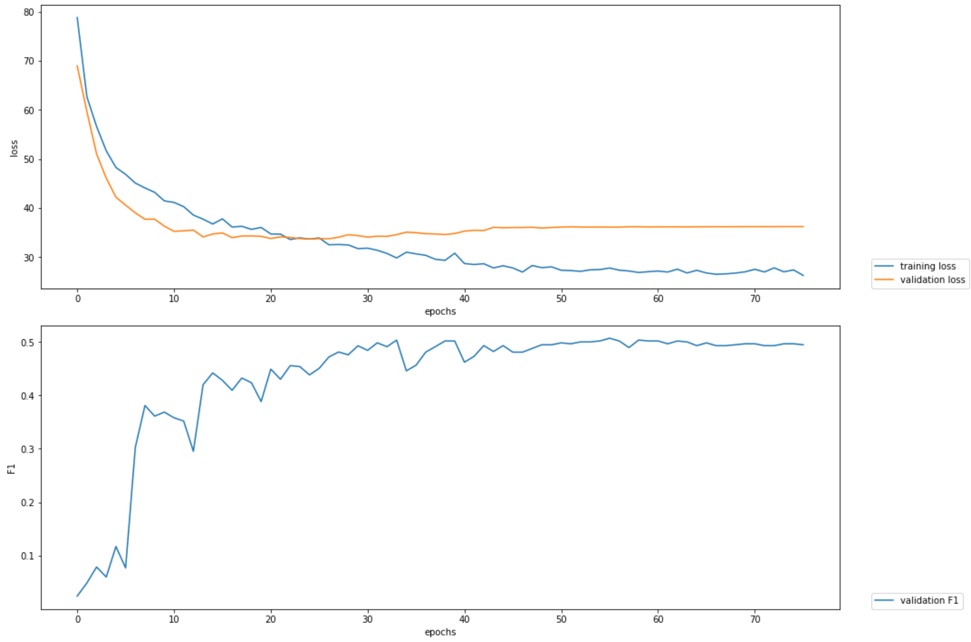
1050 s

EPOCH 76 done: loss 26.2837 - lr 0.0001953
2021-07-01 09:07:27,046 DEV : loss 36.22633361816406 - score 0.4947

KEY tp: 354 - fp: 844 - fn: 630 - precision: 0.2955 - recall: 0.3598 - f1-score: 0.3245

KEY tp: 73 - fp: 64 - fn: 78 - precision: 0.5328 - recall: 0.4834 - f1-score: 0.5069
tensor(35.2536, device='cuda:0')

KEY tp: 881 - fp: 550 - fn: 396 - precision: 0.6157 - recall: 0.6899 - f1-score: 0.6507
tensor(18.9106, device='cuda:0')



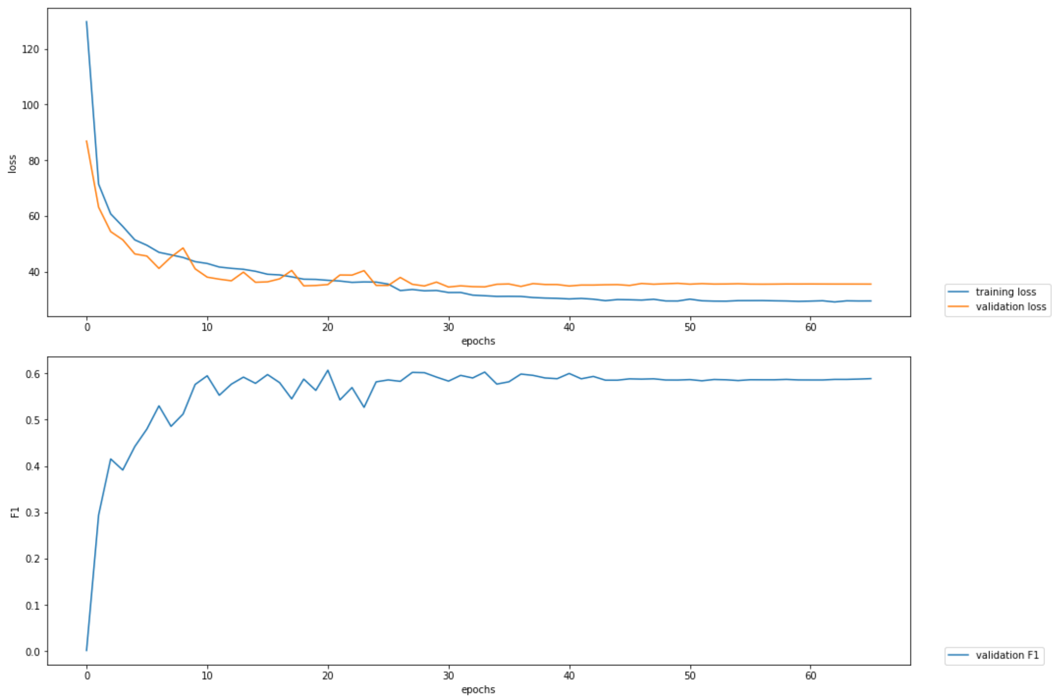
SE2017

2000 s
EPOCH 66 done: loss 29.4983 - lr 0.0001953
2021-06-30 15:48:43,829 DEV : loss 35.53535842895508 - score 0.5885

KEY tp: 1004 - fp: 934 - fn: 748 - precision: 0.5181 - recall: 0.5731 - f1-score: 0.5442

KEY tp: 616 - fp: 438 - fn: 361 - precision: 0.5844 - recall: 0.6305 - f1-score: 0.6066
tensor(34.1178, device='cuda:0')

KEY tp: 3183 - fp: 2557 - fn: 2286 - precision: 0.5545 - recall: 0.5820 - f1-score: 0.5679
tensor(30.6392, device='cuda:0')

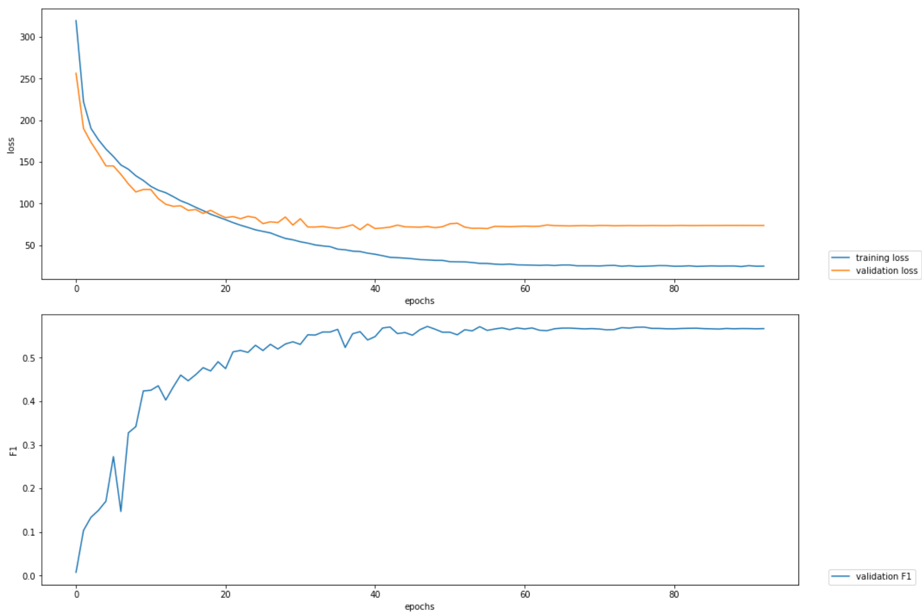


3050 s
EPOCH 93 done: loss 24.9769 - lr 0.0001953
2021-07-01 09:40:36,020 DEV : loss 73.77323913574219 - score 0.5667

KEY	tp: 1024 - fp: 1254 - fn: 728 - precision: 0.4495 - recall: 0.5845 - f1-score: 0.5082

KEY	tp: 630 - fp: 597 - fn: 347 - precision: 0.5134 - recall: 0.6448 - f1-score: 0.5717
tensor(70.5837, device='cuda:0')

KEY	tp: 4541 - fp: 1413 - fn: 928 - precision: 0.7627 - recall: 0.8303 - f1-score: 0.7951
tensor(17.2368, device='cuda:0')



www

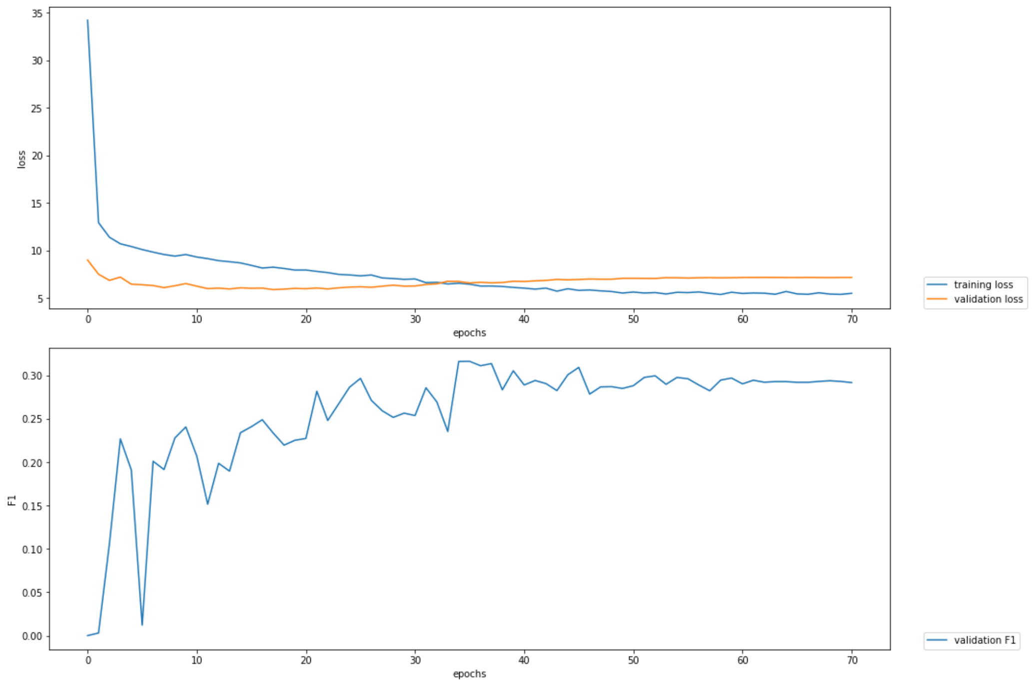
3900 s

EPOCH 71 done: loss 5.4999 - lr 0.0001953
2021-06-30 16:17:32,507 DEV : loss 7.164697170257568 - score 0.292

KEY tp: 171 - fp: 286 - fn: 443 - precision: 0.3742 - recall: 0.2785 - f1-score: 0.3193

KEY tp: 176 - fp: 275 - fn: 485 - precision: 0.3902 - recall: 0.2663 - f1-score: 0.3165
tensor(6.8917, device='cuda:0')

KEY tp: 1114 - fp: 457 - fn: 751 - precision: 0.7091 - recall: 0.5973 - f1-score: 0.6484
tensor(4.4545, device='cuda:0')

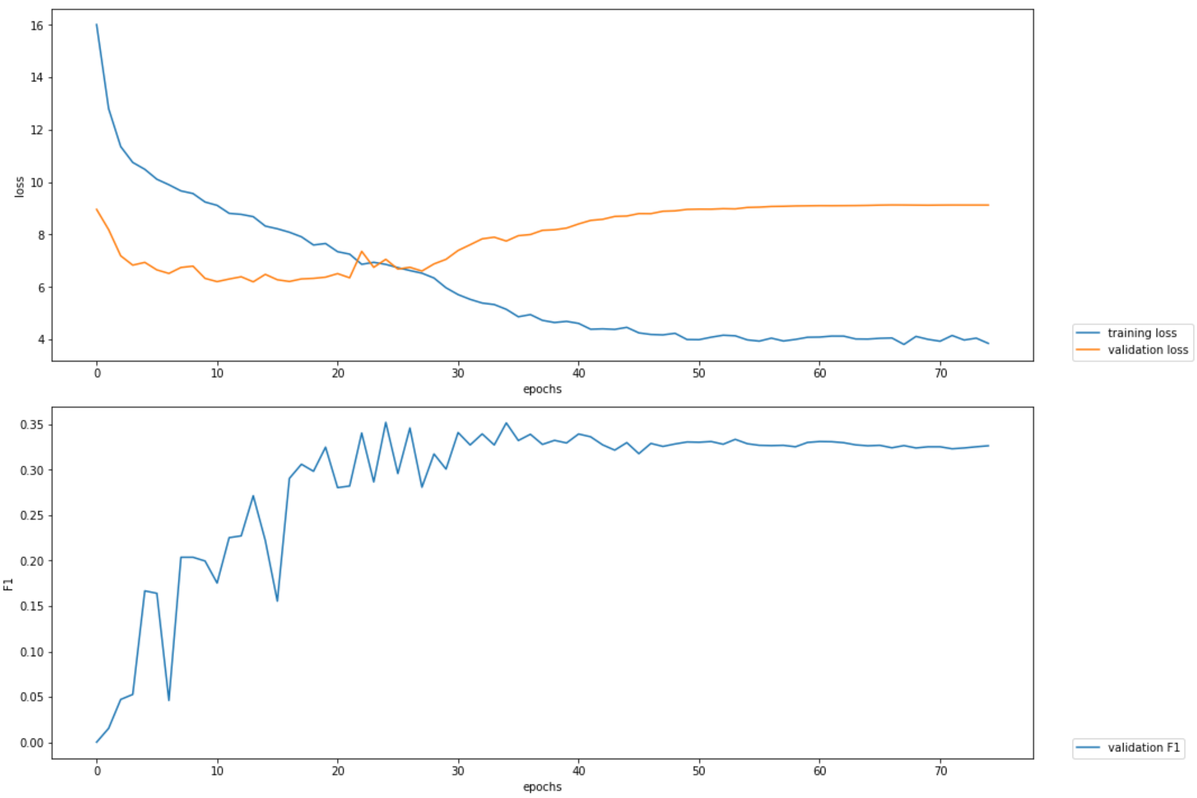


4080 s
EPOCH 75 done: loss 3.8476 - lr 0.0001953
2021-07-01 09:41:49,685 DEV : loss 9.123988151550293 - score 0.3265

KEY tp: 226 - fp: 492 - fn: 388 - precision: 0.3148 - recall: 0.3681 - f1-score: 0.3393

KEY tp: 242 - fp: 471 - fn: 419 - precision: 0.3394 - recall: 0.3661 - f1-score: 0.3523
tensor(7.3639, device='cuda:0')

KEY tp: 1358 - fp: 824 - fn: 507 - precision: 0.6224 - recall: 0.7282 - f1-score: 0.6711
tensor(4.8298, device='cuda:0')

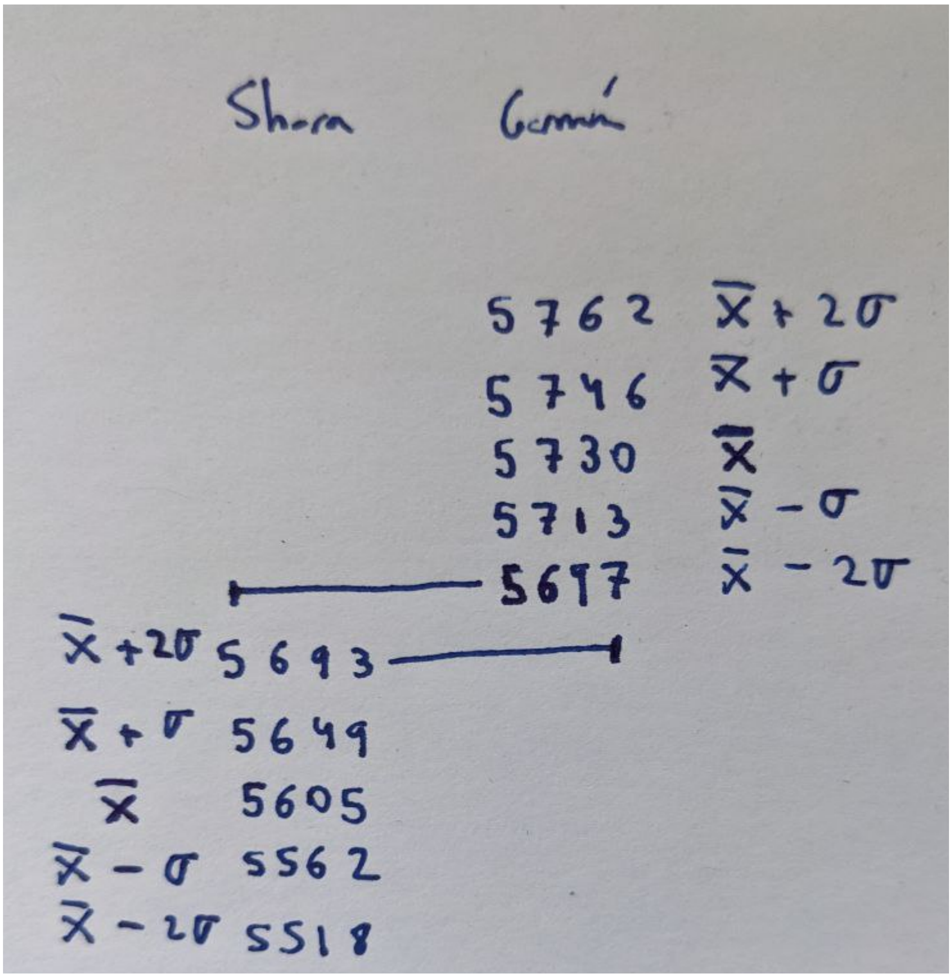

loss
epochs
training loss
validation loss
F1
epochs
validation F1

Estudio de los resultados

	Prueba_A	Prueba_B
	0,5661	0,5739
	0,5608	0,5748
	0,5661	0,5704
	0,5626	0,5747
	0,559	0,5732
	0,5538	0,5740
	0,5591	0,5754
	0,5664	0,5727
	0,5599	0,5693
	0,5616	0,5741
	0,5555	0,5730
	0,5544	0,5725
	0,5588	0,5713
	0,5679	0,5732
	0,5568	0,5725
MEAN	0,560587	0,573
STD	0,004368	0,001611
M + S	0,564954	0,574611
M - S	0,556219	0,571389

	Prueba_A	Prueba_B
		0,5739
	0,5661	
	0,5608	0,5748
	0,5661	0,5704
	0,5626	0,5747
	0,559	0,5732
	0,5538	0,5740
	0,5591	0,5754
	0,5664	0,5727
	0,5599	0,5693
	0,5616	0,5741
	0,5555	0,5730
	0,5544	0,5725
	0,5588	0,5713
	0,5679	0,5732
	0,5568	0,5725

MEAN	0,560587	0,573
STD	0,004368	0,001611
M + 2S	0,569322	0,576222
M - 2S	0,551851	0,569778


5762 $\bar{X}+2\sigma$
5746 $\bar{X}+\sigma$
5730 $\bar{X}$
5713 $\bar{X}-\sigma$
5697 $\bar{X}-2\sigma$
$\bar{X}+2\sigma$ 5693
$\bar{X}+\sigma$ 5649
$\bar{X}$ 5605
$\bar{X}-\sigma$ 5562
$\bar{X}-2\sigma$ 5518

Recursos:

Yake en Cacic 0.196

we conclude that good keyphrases
of a document should satisfy the following
properties:
1. Understandable. The keyphrases are understandable
to people. This indicates the
extracted keyphrases should be grammatical.
For example, “machine learning” is a grammatical
phrase, but “machine learned” is not.
2. Relevant. The keyphrases are semantically
relevant with the document theme. For example,
for a document about “machine learning”,
we want the keyphrases all about this
theme.

3. Good coverage. The keyphrases should
257
cover the whole document well. Suppose
we have a document describing “Beijing”
from various aspects of “location”,
“atmosphere” and “culture”, the extracted
keyphrases should cover all the three aspects,
instead of just a partial subset of them.

As described
in (Wan and Xiao, 2008b), when two annotators
were asked to label keyphrases on 308 documents,
the Kappa statistic for measuring interagreement
among them was only 0.70.

There are three approaches to address keyphrase extraction:
(i) traditional two-step ranking method, (ii) sequence labeling
and (iii) generation using neural networks.

Word embeddings have already shown promising
results in the process of keyphrase extraction
from scientific articles (Wang et al., 2015, 2014).
However,Wang et al. did not use domain-specific
word embeddings and had suggested that training
them might lead to improvements. This motivated
us to experiment with do

Rui Wang, Wei Liu, and Chris McDonald. 2014.
Corpus-independent generic keyphrase extraction
using word embedding vectors. In Software Engineering
Research Conference. volume 39.
RuiWang,Wei Liu, and Chris McDonald. 2015. Using
word embeddings to enhance keyword identification
for scientific publications. In Australasian Database
Conference. Springer, pages 257–268.

INSPEC

We choose the publicly available dataset, Hulth20031, a collection of 2,000 abstracts of Computer Science journal articles, each of approximately 100-150 words, extracted from Inspec. Each document has two human assigned keyword lists - author assigned keywords (restricted to a certain number) and reader assigned ones (freely assigning), where 78% readers' assigned ones appear in the texts, in contrast to only 25% for authors' assigned keywords. The total number of keywords assigned by readers is 19,276, and 8,945 by authors, in which only 1,207 are in common. Moreover, we find that author assigned keywords tend to be more abstractive, often describing board fields, whereas the reader assigned ones tend to be more extractive.

Ordenados:

Temas:

KEYWORDS

2003_improved

supervised machine learning

adding linguistic knowledge
to the representation (such as syntactic
features), rather than relying only on
statistics (such as term frequency and ngrams),
a better result is obtained

NP-chunks gives a better precision
than n-grams, and by adding the POS
tag(s) assigned to the term as a feature

2009_clustering

unsupervised method
for keyphrase extraction

method finds exemplar terms by leveraging
clustering techniques, which guarantees
the document to be semantically covered
by these exemplar terms

keyphrases are extracted from the document
using the exemplar terms

outperforms sate-of-the-art graphbased
ranking methods (TextRank) by
9.5% in F1-measure.

Keyphrases in articles of journals and books
are usually assigned by authors.
most articles on the web usually do not have
human-assigned keyphrases. Therefore, automatic
keyphrase extraction is an important research task.

The supervised approach (Turney, 1999) regards keyphrase extraction as a classification task.

For the supervised keyphrase extraction
approach, a document set with human-assigned
keyphrases is required as training set.
human labelling is time-consuming.

the supervised methods
need manually annotated training set, which may
sometimes not be practical, especially in the web
scenario.

A straightforward method for keyphrase extraction
is to select keyphrases according to frequency
criteria.

different learning
algorithm, Naive Bayes method, is applied in
(Frank et al., 1999) with improved results on the
same data used in (Turney, 1999). Hulth (Hulth,
2003; Hulth, 2004) adds more linguistic knowledge,
such as syntactic features, to enrich term
representation, which significantly improves the
performance.

Starting with TextRank (Mihalcea and Tarau,
2004), graph-based ranking methods are becoming
the most widely used unsupervised approach
for keyphrase extraction.

We should also note that, although the precision
and recall of most current keyphrase extractors
are still much lower compared to other NLPtasks,
it does not indicate the performance is poor
because even different annotators may assign different
keyphrases to the same document.

2010_HUMB

Analysis of the structure of the article.
2. Selection of candidate terms.
3. Calculation of features.
4. Application of a ML model for evaluating
each candidate term independently.
5. Final re-ranking for capturing relationships
between the term candidates.

Utiliza varios modelos de ML (Decision tree, Multi-Layer perceptron, Support Vector Machine)

Key terms (or keyphrases or keywords) are metadata providing general information about the content of a document. Their selection by authors or readers is, to a large extent, subjective which makes automatic extraction difficult.

2013_TopicRank

graph-based ranking model is applied to assign a significance score to each topic. Keyphrases are then generated by selecting a candidate from each of the topranked topics.

This new method
is an improvement of the TextRank method applied to keyphrase extraction (Mihalcea and Tarau, 2004). In the TextRank method, a document is represented by a graph where words are vertices and edges represent co-occurrence relations. A graph-based ranking model derived from PageRank (Brin and Page, 1998) is then used to assign a significance score to each word.

ranking topics instead of words
is a more straightforward way to identify the set of keyphrases that covers the main topics of a document.

Although scientific articles usually provide them, most of the documents have no associated keyphrases. Therefore, the problem of automatically assigning keyphrases to documents is an active field of research.

Supervised methods recast
keyphrase extraction as a binary classification task (Witten et al., 1999),

While supervised
approaches have generally proven to be more successful,

the need for training data and the bias towards
the domain on which they are trained remain
two critical issues

whereas unsupervised
methods apply different kinds of techniques
such as language modeling (Tomokiyo and Hurst,
2003), clustering (Liu et al., 2009) or graph-based
ranking (Mihalcea and Tarau, 2004).

2015_SGRank

Híbrido estadístico - gráfico. Se realizan n-grams, se puntúan según una versión de tf-idf, los mejores candidatos se analizan según otras heurísticas estadísticas (posición de la primera ocurrencia, longitud del término) se le aplica un algoritmo basado en grafo que devuelve el rank. **Valores Pre y Rec aprox*

hybrid statistical-graphical
algorithm that capitalizes on the heuristics of
both families of algorithms

1) Construct a keyphrase
extraction algorithm based on optimal statistical
features and 2) Combine it with a graph-based
algorithm for further improvements. The advantage
of graph-based methods is that they take into
account term co-occurrence patterns that are not
generally utilized by statistical methods which take
a bag of n-grams approach to document
representation.

Two major families of unsupervised
keyphrase extraction algorithms may be
characterized as statistical and graph-based

keyphrase
extraction can be used to facilitate the automatic
construction of concept maps (Leake et al., 2003)conceptMap
or ontologies (Fortuna et al., 2006)ontologias

Keyphrase
extraction is also used in content-based
recommender systems which help users in
discovering information relevant to their
previously expressed interests (Lops et al., 2011).

There are two prominent families of unsupervised keyphrase extraction algorithms. The older of these two is clustered around the tf-idf term weighting metric where word statistics such as frequency of occurrence in the document or rareness in the corpus are used to distinguish potential keyphrases. The more recently developed of the two families has been built on the

foundation of the TextRank algorithm (Mihalcea & Tarau, 2004). In algorithms of this family a graphical representation of the text is constructed with words as nodes and edges reflecting cooccurrence relations. This graph is then used to run node ranking algorithms such as PageRank (Page et al., 1999) that assign weights to the node-words reflecting their semantic importance to the text.

term frequency-inverse document frequency (tf-idf) term weighting function (Salton et al, 1975). Given a corpus of documents the tf-idf weight of term t in document d is mathematically expressed as tf-idf(t,d)=tf(t,d)*idf(t) where tf(t,d) is the frequency of term t in document d and idf(t)=log(N/df(t)) where N is the total number of documents in the corpus and df(t) is the number of documents in the corpus that contain term t (Jones, 1972). The term frequency heuristic is based on the intuition that terms which occur more often in a document are more likely to be important to its meaning. The idf function captures the rareness heuristic, that is, words which occur in many documents in the corpus are unlikely to be important to the meaning of any specific one.

2016_keyphrase

addressing keyphrase extraction as a *sequential labeling* task

Uso de una capa CRF. Método de sequence modeling.No pone el uso de 10 keyphrases

Several applications in data mining and NLP in- volve predicting an output sequence of labels y = <y_1 ...y_N> given an input sequence of to- kens: t = <t_1 ...t_N> where each input posi- tion i : 1 . . . N is modeled by vectors of features <x_1 . . . x_N > (Sarawagi, 2005). For example, given a text sequence (of words), Named-Entity Recogni- tion (NER) involves predicting labels for each word from the set {*person, organization, location, other*} using various dictionary, linguistic, and surface- pattern features at each position.

Taggers involving complex, interdependent features are often trained using discriminative learning algorithms such as **Conditional Random Fields (or CRFs)**. CRFs com- prise state-of-the-art models for several sequence tagging tasks and hence we use them for learning a keyphrase tagger (Sutton and McCallum, 2012).

Mas info en el paper

2017_MIKE

multidimensional information to improve the keyphrase extraction performance

Modelo gráfico. Resuelve el problema a modo de problema de optimización..

The supervised approaches generally treat the keyphrase extraction as a binary classification task, in which a learning model is trained on the features of labelled keyphrases to determine whether a candidate phrase is a keyphrase.

Current state-of-the-art supervised models include Na ̈ive Bayes [8, 9, 13], Decision Trees [34, 37], Maximum En- tropy [45], Random Forest [1], Multi-layer Perceptron [28], Support Vector Machines (SVM) [28], Conditional Random Fields [15], Inte- ger Linear Program [7, 10], Deep Recurrent Neural Networks [46] etc.

In contrast, unsupervised approaches directly treat keyphrase extraction as a ranking problem, scoring each candidate using dif- ferent kinds of techniques such as language modelling [36], cluster- ing [27] or graph-based ranking [12, 14, 26, 31, 35, 38]. In particular, the graph-based algorithms are widely used in the unsupervised scenario. These approaches first build a word graph (as illustrated in Fig. 1(2)) in which each node denotes a candidate word and each edge/link prepresents a relevant relation (e.g., co-occurrence relation) between candidates within a document. Subsequently, var- ious centrality measures [6] or random walk techniques (e.g., Page- Rank [32]) are used on the word graph to rank candidate words.

2018_bidirectional

or non-NLP tasks such as social network analysis [23] or user modeling [24]. Automatic Keyphrase Extraction

Since no algorithm stands out as the

\best" ML algorithm, often authors test many techniques in a single experiment, and then they choose as best ML algorithm the best performing one [2,9] and/or even the least computationally expensive one [19].

a supervised approach, was the best performing algorithm in the SEMVAL 2010 Keyphrase Extraction Task [16].

In the last years, most attention is devoted to the features used in these supervised algorithms. The numbers of features used can range from just two [32] to more than 20 [9]. These features can be divided in categories identied

with di
erent kinds of knowledge they encode into the model:

{ statistical knowledge. number of appearances of the KP in the document, TF-IDF, number of sentences containing the KP, etc.;

{ positional knowledge. position of the rst occurrence of the KP in the document, position of the last occurrence, appearance in the title, appearance in specic sections (abstract, conclusions), etc.;

{ linguistic knowledge: part-of-speech tags of the KP [14], anaphoras pointing to the KP [2], etc.;

external knowledge: presence of the KP as a page on Wikipedia [6] or in specialized domain ontologies [19], etc.

In recent years, Deep Learning techniques have shown impressive results in many Natural Language Processing tasks, e.g., Named Entity Recognition, Automatic Summarization, Question Answering, and so on [18,27,25,29]. In Named Entity Recognition, for example, researchers have proposed several Neural Network Architectures

In [33], the

authors present an approach based on Recurrent Neural Networks, specically designed for a particular domain, i.e., Twitter data.

Bidirectional Long Short-Term

Memory RNN (hence Bi-LSTM), which is able to exploit previous and future context of a given word. Our system, since it does not require specic features carefully optimized for a specic domain, can be applied to a wide range of scenarios.

As in the techniques used for Named Entity Recognition, the three neurons are mapped to three possible output classes: NO KP, BEGIN KP, INSIDE KP, which respectively mark tokens that are not keyphrases, the rst token of a keyphrase, and the other tokens of a keyphrase.

For example, if our input sentence is \We train a neural network using Keras", and the keyphrases in that sentence are \neural network" and \Keras", the tokens' classes will be We/NO KP train/NO KP a/NO KP neural/BEGIN KP network/INSIDE KP using/NO KP Keras/BEGIN KP'.

The input layer of our model is a vector representation of the individual words contained in input document. Several recent studies [5,22] showed that such representations, called word embeddings, are able to represent the semantics of words better than an \one hot" encoding word representation, when trained on large corpus. **However, the datasets for AKE are relatively small, therefore it is dicult to train word embeddings to capture the word semantics. Hence,**

we adopt Stanford's GloVe Embeddings, which are trained on 6 billion words extracted from Wikipedia and Web texts [26].

2018_Key2Vec

Utiliza embeddings específicos del dominio para mejorar los resultados. Utiliza un método basado en grafos parecido a textrank.

2019_Bi-LSTM-CRF

keyphrase extraction problem as sequence labeling and propose a model that jointly exploits the complementary strengths of Conditional Random Fields that capture label dependencies through a transition parameter matrix consisting of the transition probabilities from one label to the neighboring label, and Bidirectional Long Short Term Memory networks that capture hidden semantics in text through the long distance dependencies.

More recently, Gollapalli et al. [13] formulated keyphrase extraction as sequence labeling and showed that using linear-chain Conditional Random Fields can improve the performance over baseline models for this task. However, the approach in [13] does not explicitly take into account the long-term dependencies and semantic relationships hidden in text.

focus on keyphrase extraction, i.e., extracting only words that are present in text, and not keyphrase generation, which outputs words that may or may not be present in text

Sequence labeling models for keyphrase extraction have shown promising results in recent studies [4, 13, 44]. For example, Gollapalli et al. [13] trained Conditional Random Fields (CRFs) to extract keyphrases from scholarly documents, using features such as tf-idf

and POS tags to predict a label for each token position in a document **as being a keyphrase token (KP) or not (Non-KP)**. This CRF model is able to capture the dependencies in previous and future tags in the label sequence, however, the semantic dependencies in the input sequence, i.e., the text, are not incorporated.

INFO SOBRE Bi-LSTM

2019_Keyphrase

Dhruva Sahrawat

2019_Unsupervised

BERT + BiLSTM-CRF + document-level attention (extrae la información más relevante de cada oración)

2020_DAKE

Utiliza BERT y BiLSTM-CRF. Bien explicado cada capa y fase

2020_Keyphrase

There are three main approaches: (i) traditional twostep ranking method, (ii) sequence labeling method and (iii) generation-based neural network method. A straightforward

way of keyphrase extraction is to decompose this
task into two steps: candidate phrases generation and candidate
phrases scoring (Witten et al. 2005; Medelyan, Frank,
A weighted ranking algorithm is proposed.
weighted ranking algorithm
ranking algorithm
Figure 1: The overlap phenomenon in keyphrases.
and Witten 2009). In the first step, models generate a list
of candidate phrases using the N-grams or phrases with certain
part-of-speech patterns. In the second step, candidate
phrases are scored by its probability of being a keyphrase in
the given document. These two-stage ranking-based methods
treat candidate phrases individually, which makes it
almost impossible to capture the contextual information
of candidates and interactions between different phrases.
Further, existing two-stage methods (Witten et al. 2005;
Medelyan, Frank, andWitten 2009) are based on feature engineering,
which is labor intensive and domain dependent.
Another intuitive approach is to regard keyphrase extraction
as a sequence labeling task (Zhang et al. 2016). However, sequence
labeling approach can hardly tackle keyphrase with
overlapping words. As shown in Figure 1, “weighted ranking
algorithm” and “ranking algorithm” are both keyphrases
that provide semantic information of different granularity.
Unfortunately, sequence labeling methods are not able to
extract both of them at the same time. With the development

of deep learning, especially sequence-to-sequence methods, generation-based methods (Meng et al. 2017; Chen et al. 2018) have attracted much attention. Admittedly, generation-based approach is capable of dealing with the overlap keyphrases without much labor-intensive feature engineering, but such an approach has two shortcomings. Firstly, generation method can not utilize context information effectively and capture the interaction of phrases. Secondly, generation method produces tokens of keyphrases continuously which suffers the problem of not utilizing phrases-level information. In other words, tokens in a phrase can only make sense when they appear as a whole sequence. Therefore, a desirable solution should be able to capture the information within word sequences and take advantage of this span-based information when predicting the keyphrases.

The proposed
SKE model first extracts candidate phrases using the certain part-of-speech patterns (Le, Nguyen, and Shimazu 2016) and records the beginning and ending positions of each candidate phrase as spans. After that, Bert (Devlin et al. 2019) or recurrent neural network based on word vectors is used to represent the high-level concept of phrases. We call the high-level representation span-based representation. Afterward, bidirectional recurrent neural network (i.e., LSTM and GRU) is used to capture the interaction of spanbased

representation to get higher-level phrase representation.

After getting the phrase representation, we are able to

use them to classify the candidate phrase.

BUEN RELATED WORK

2020_WEKE

the supervised methods treat the
keyphrase extraction as a binary classification task, in which a classifier is trained
on the features of labeled keyphrases to determine whether a candidate phrase
is a keyphrase [5].

In unsupervised methods, on the other hand, various measures,
such as tf-idf (term frequency-inverse document frequency), graph-based
ranking scores (e.g., degree centrality and PageRank score), are used to score
individual candidate words that are later summed up to obtain total scores for
phrases [3,4,8,11,21,22].

The PageRank-based models (i.e., random-walk models) are widely used in
the unsupervised scenario and considered as the current state of the arts. These
models first build a word graph in which each node denotes a candidate word and
each edge represents a co-occurrence relation between words within a document.
Random walk techniques are subsequently used on the word graph to rank words.
Since TextRank [11] firstly computed ranking scores of candidate words using
PageRank algorithm [13] on the word graph, many TextRank-based extensions

have been proposed, aiming at integrating various types of information into modified PageRank models to improve the performance of keyphrase extraction. For example, TopicalPageRank (TPR) [8], CiteTextRank [4] and PositionRank [3] integrate topical information, citation context and position information into the PageRank-based framework, respectively.

Although remarkable efforts have been made on PageRank-based framework for keyphrase extraction, these traditional models cannot utilize the local context information of the word graph, which can be considered as the local collocation pattern between words at the semantic level or as the semantic similarity between words. The reason for this is that the PageRank-based ranking model is a way to decide the importance of a node within a given graph, by taking into account global structural information recursively computed from the entire graph, rather than relying only on local structural information [11].

2020_YAKE

the difficulties of defining relevance and the language diversity itself (e.g., texts of different languages, sizes, and domains) are among the most im- portant challenges. Other issues involve the problem posed by absent keywords, 3 the exact match restriction, 4 and the high number of candidate keywords that can be generated from a single text. 5

In an era of vast but often unlabeled collections, this can be a great advantage over other approaches in cases where access to a training corpus is either limited or restricted.

BUEN RESUMEN UNSUPER Y SUPER

Keyword extraction: Issues and methods

ph.

Textual keyword extraction and summarization: State-of-the-art

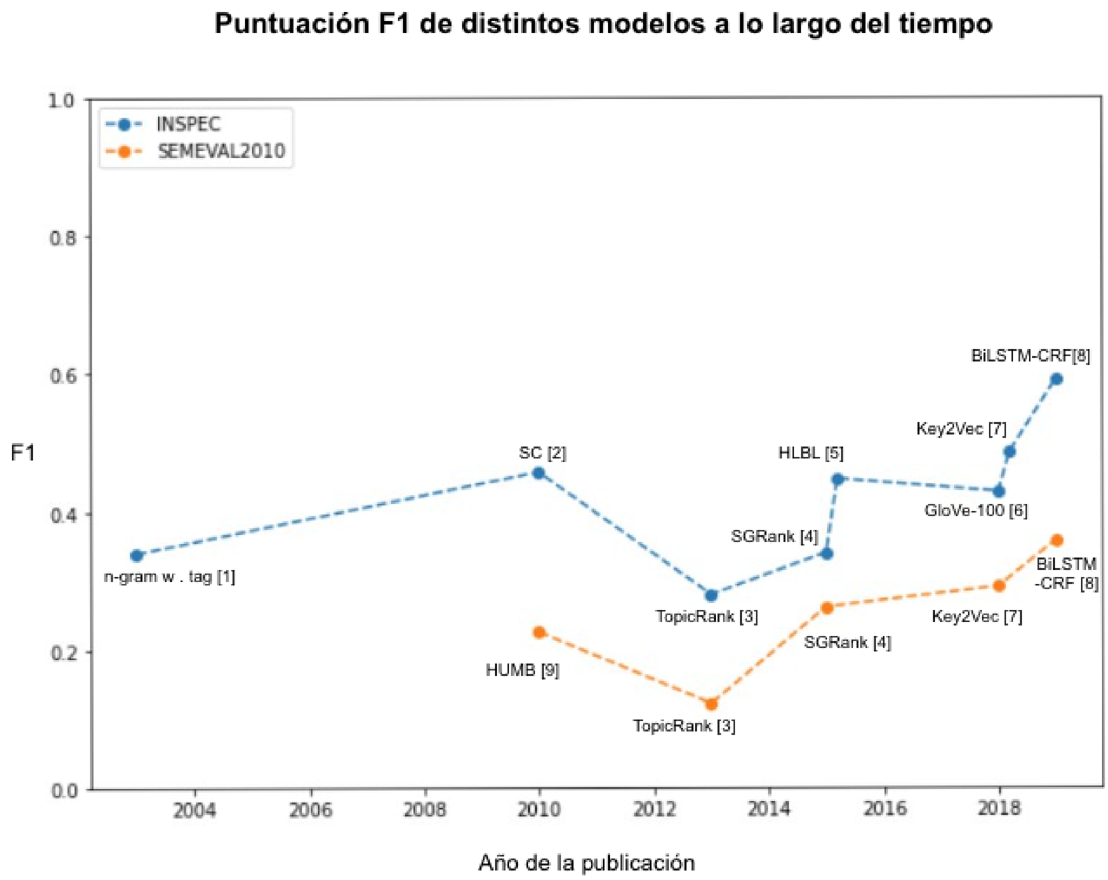
BUENA EXPLICACIÓN DE LAS MÉTRICAS

BERT (*Bidirectional Encoder Representations from Transformers*) o **Representación de Codificador Bidireccional de Transformadores** es una técnica basada en redes neuronales para el pre-entrenamiento del procesamiento del lenguaje natural (PLN) desarrollada por Google.[1] BERT fue creado y publicado en 2018 por Jacob Devlin y sus colegas de Google.[2] Google está aprovechando BERT para comprender mejor las búsquedas de los usuarios. [3]

El modelo BERT original se creó usando dos corpus de lengua inglesa: BookCorpus y Wikipedia en inglés.

BERT (Bidirectional Encoder Representations from Transformers): este modelo de representación de lenguaje creado por Google AI Language fue diseñado para pre-entrenar representaciones bidireccionales de texto no etiquetado teniendo en cuenta el contexto en ambas direcciones [27]. Los modelos de lenguaje desarrollados hasta el momento tenían la limitación de ser unidireccionales, lo cual empeora las capacidades para llevar a cabo tareas de PLN, en las que es crucial incorporar el contexto en ambas direcciones. Para ello, BERT emplea un modelo de lenguaje enmascarado. Este tipo de modelos realiza un enmascarado en determinados tokens1 de entrada de manera aleatoria, con el objetivo de predecir la identificación del vocabulario original de la palabra enmascarada basándose únicamente en su contexto. La estructura

Recursos antiguos



[5] *Using word embeddings to enhance keyword identification for scientific publications* - Wang R, Liu W, McDonald C (2015)

Procesamiento del lenguaje natural.

El Procesamiento del Lenguaje Natural (a partir de ahora NLP, de sus siglas en inglés, Natural Language Processing) es una rama de la informática, más en concreto de la inteligencia artificial que tiene como objetivo dar la capacidad a los ordenadores de analizar y comprender tanto el habla como los textos escritos necesarios para la interacción humano-máquina usando el lenguaje natural. El NLP combina conceptos de distintas áreas del conocimiento como son la informática, la lingüística, la psicología, las matemáticas o la ciencia cognitiva.

La lingüística es el estudio científico del lenguaje, incluyendo su gramática, semántica y fonética. La lingüística clásica consiste en idear y evaluar reglas del lenguaje. Estas reglas y estructuras pueden ser modeladas de manera matemática, campo denominado como la lingüística matemática que se centra en el uso de formalismos matemáticos discretos y la teoría para el lenguaje natural. De esta forma nace la lingüística computacional, que utiliza las herramientas que proporciona la informática para el estudio de una manera más eficiente y con el objetivo de que las máquinas sean capaces de comprender el lenguaje humano.

La tendencia con el NLP es la combinación de lingüística computacional (modelado basado en reglas del lenguaje humano) con modelos estadísticos, de aprendizaje automático o aprendizaje profundo.

Sus inicios se remontan a los años 50, donde con la proposición de Alan Turing del test de Turing se trató de abordar el problema del procesamiento del lenguaje natural, durante estos años, el objetivo fue la creación de gramáticas que un ordenador fuese capaz de entender y traducir nuestro lenguaje a estas nuevas gramáticas. No obstante pese a los intentos del National Research Council de estados unidos, con la creación de ALPAC investigando en la problemática, no se consiguió ningún progreso, haciendo el proceso de traducción de una máquina más costoso que por humanos, y no consiguiendo que un ordenador mantuviese una conversación básica. No fue hasta varios años más tarde, a finales de los años 80, que gracias al aumento de la potencia computacional y los orígenes del aprendizaje automático, se retomó esta disciplina. Más en concreto con el uso de redes neuronales, que en 2001, el equipo de Yoshio Bengio, propuso el primer modelo de lenguaje neuronal usando una red neuronal feed-forward (tipo de arquitectura de una red neuronal que tiene el flujo de datos de manera unidireccional sin hacer ciclos). A partir de este momento, con la gran investigación e inversión que se está dando al estudio de las redes neuronales, los resultados del NLP van mejorando de manera simultánea con los avances y las creaciones de estos nuevos modelos hasta la actualidad, donde tanto el NLP como las redes neuronales prometen tener un gran futuro.

Aplicaciones del NLP

Reconocimiento del habla

Etiquetado gramatical

Desambiguación lingüística

Análisis de sentimiento

Traducción automática

Recuperación de la información

Extracción de la información (Hilar con extracción de términos)

Extracción de términos: