GeoAI LMR: TETIS Text mining

remy.decoupes

December 2022

Abstract

In the context of the GeoAI challenge organized by the ITU, we present, in this report, our work. The challenge, dedicated to detect Location in twett content, was organized by ITU[1] and QCRI[2]

1 Introduction

2 Model selection

3 Model fine-tuning

3.0.1 Hyper-parameters

3.1 Data Augmentation

4 Results

5 Docker submission

6 Discussion

7 conclusion

[1]International Telecommunication Union
[2]Qatar Computing Research Institute

	California Wildfires 2018	Canada Wildfires 2016	Cyclone Idai 2019	Ecuador Earthquake 2016	Greece Wildfires 2018	Hurricane Dorian 2019	Hurricane Florence 2018	Hurricane Harvey 2017	Hurricane Irma 2017	Hurricane Maria 2017	Hurricane Matthew 2016	Italy Earthquake Aug 2016	Kaikoura Earthquake 2016	Kerala Floods 2018	Maryland Floods 2018	Midwestern US Floods 2019	Pakistan Earthquake 2019	Puebla Mexico Earthquake 2017	Srilanka Floods 2017	**Average**
TETIS	0.92	0.75	0.89	0.94	0.93	0.88	0.77	0.93	0.87	0.91	0.95	0.9	0.92	0.89	0.91	0.93	0.88	0.93	0.93	0.90
baseline CRF	0.98	0.94	0.89	0.92	0.93	0.96	0.74	0.97	0.97	0.97	0.97	0.82	0.98	0.96	0.94	0.92	0.94	0.92	0.94	0.93
baseline BERT	0.98	0.93	0.92	0.95	0.92	0.97	0.78	0.98	0.96	0.97	0.97	0.87	0.98	0.96	0.93	0.94	0.95	0.93	0.93	0.94

City F1	Cnty F1	Cont F1	Country F1	District F1	Hpoi F1	Island F1	Nbhd F1	Npoi F1	Other F1	State	**Global F1**
0.76	0.76	0.7	0.92	0.24	0.18	0.77	0	0.54	0	0.89	0.82