



GTÜ BİL MUH BİL 495

BİL 495
1.İzleme

Gökhan HAS

Proje Danışmanı: Dr. Burcu YILMAZ
Kasım 2020

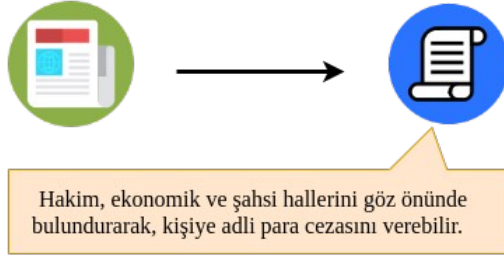


- Proje Tasarım Planı
- Veri Kümesi
- Kullanılan Model
- Parametre Analizi
- Tahmin Sonuçları
- Kaynaklar



Proje Tasarım Planı

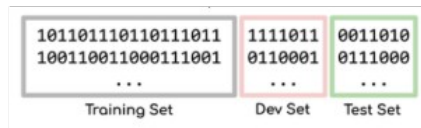
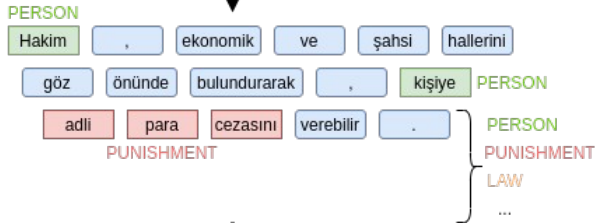
1. HAZIRLIK



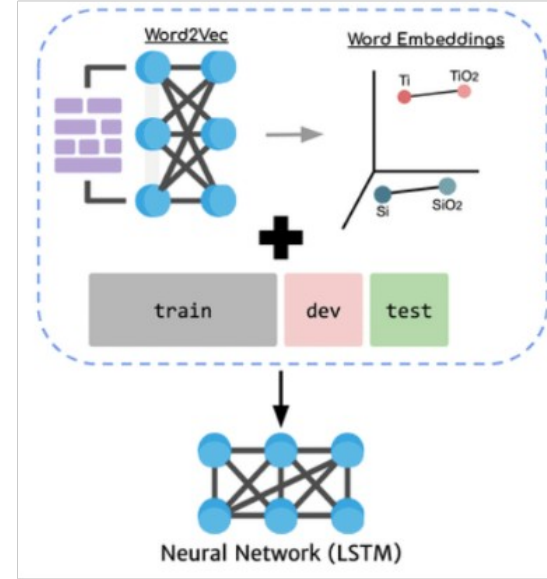
2. TOKENİZASYON



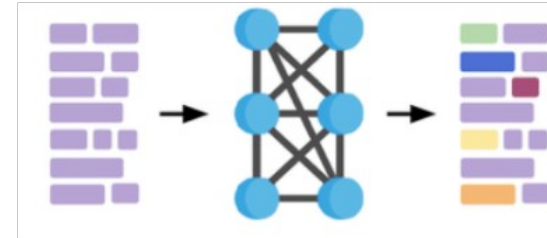
3. ETİKETLEME

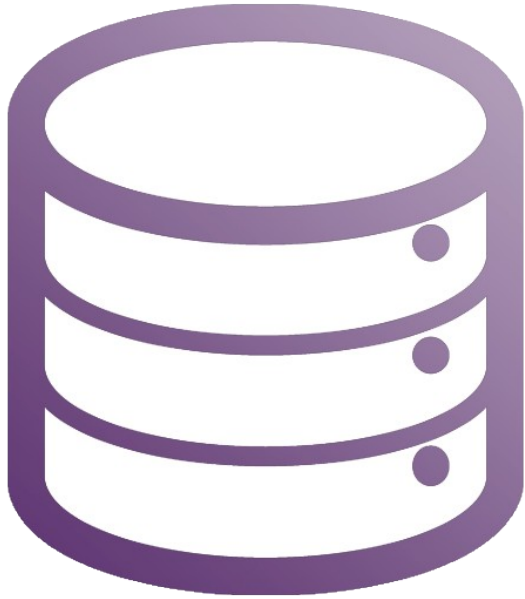


4. MODEL EĞİTİLMESİ

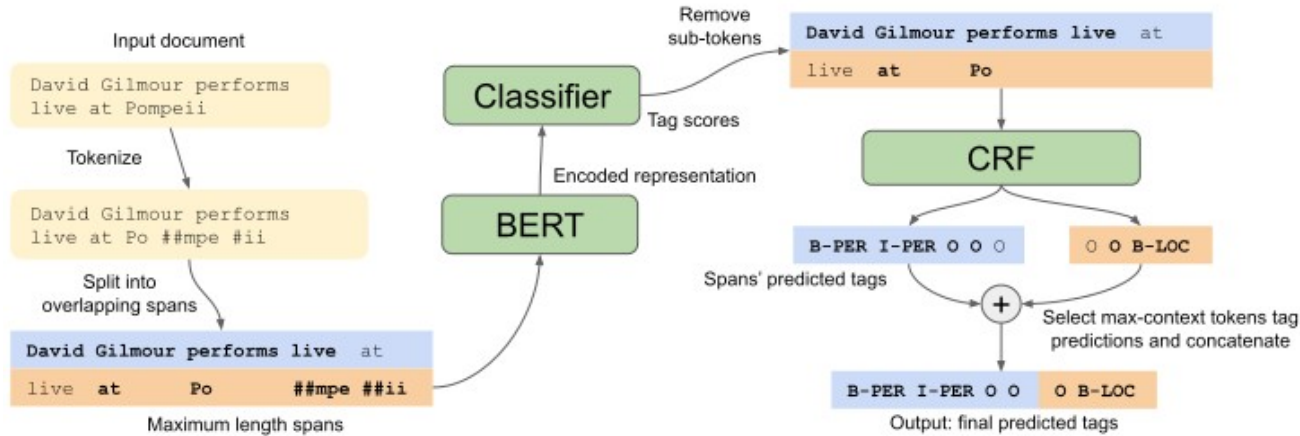


5. MODEL YARDIMIYLA SONUÇLARIN ÇIKARILMASI





Kullanılan Model - 1



HUGGING FACE

[↑ Back to all models](#)

Model: **dbmdz/bert-base-turkish-128k-cased**

pytorch

tf

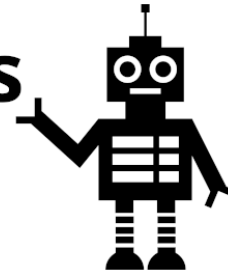
bert

tr

license:mit



Simple Transformers



BERT MODELLERİ OPTİMUM PARAMETRE DEĞERLERİ			
Learing Rate	2,00E-05	3,00E-05	5,00E-05
Batch Size	8	16	32
Train Epoch	3	4	5

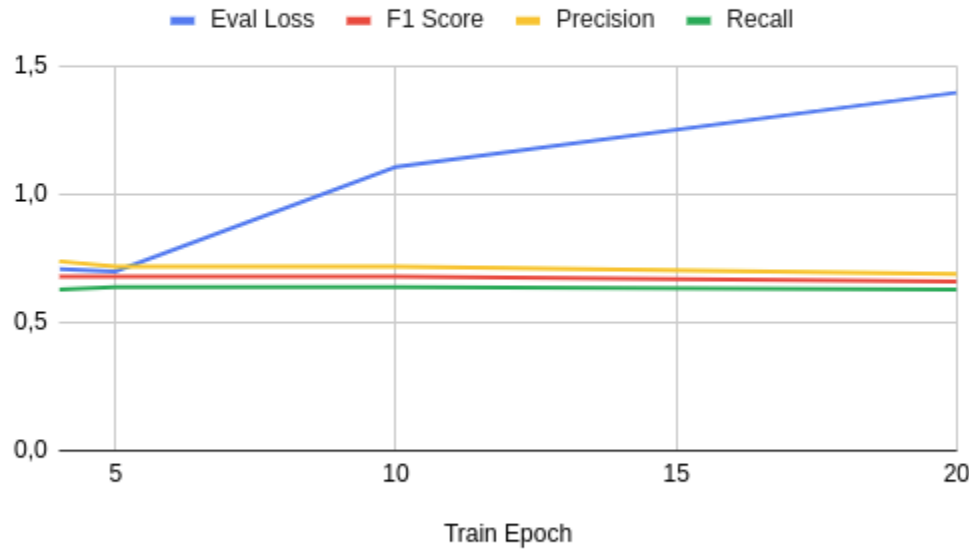
Learing Rate	
1,00E-05	0,00001
2,00E-05	0,00002
3,00E-05	0,00003
5,00E-05	0,00005



Parametre Analizi - 1

Train Epoch	Eval Loss	F1 Score	Precision	Recall
4	0,71	0,68	0,74	0,63
5	0,7	0,68	0,72	0,64
10	1,11	0,68	0,72	0,64
20	1,4	0,66	0,69	0,63

Eval Loss, F1 Score, Precision ve Recall



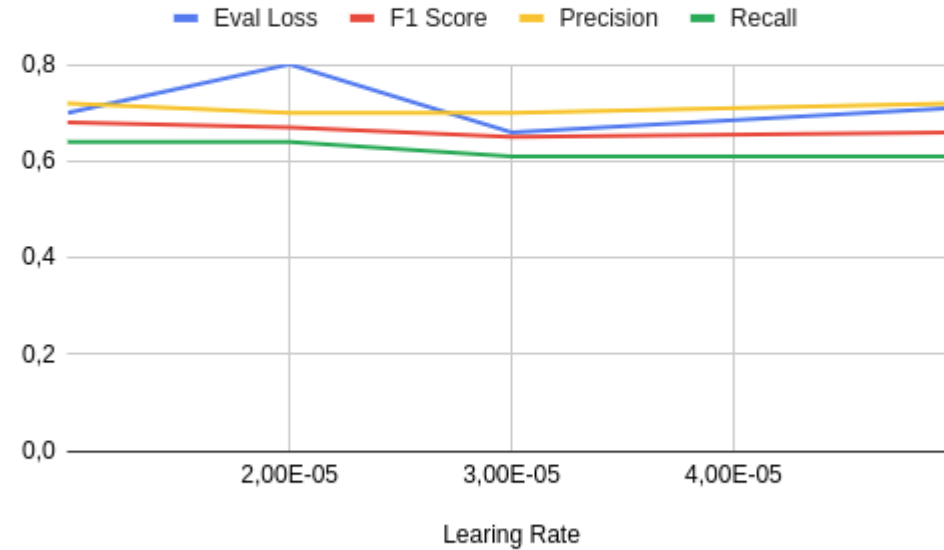
			Learning Rate	1,00E-05
			Train Batch Size	8
			Eval Batch Size	32



Parametre Analizi - 2

Learning Rate	Eval Loss	F1 Score	Precision	Recall
1,00E-05	0,7	0,68	0,72	0,64
2,00E-05	0,8	0,67	0,7	0,64
3,00E-05	0,66	0,65	0,7	0,61
5,00E-05	0,71	0,66	0,72	0,61

Eval Loss, F1 Score, Precision ve Recall



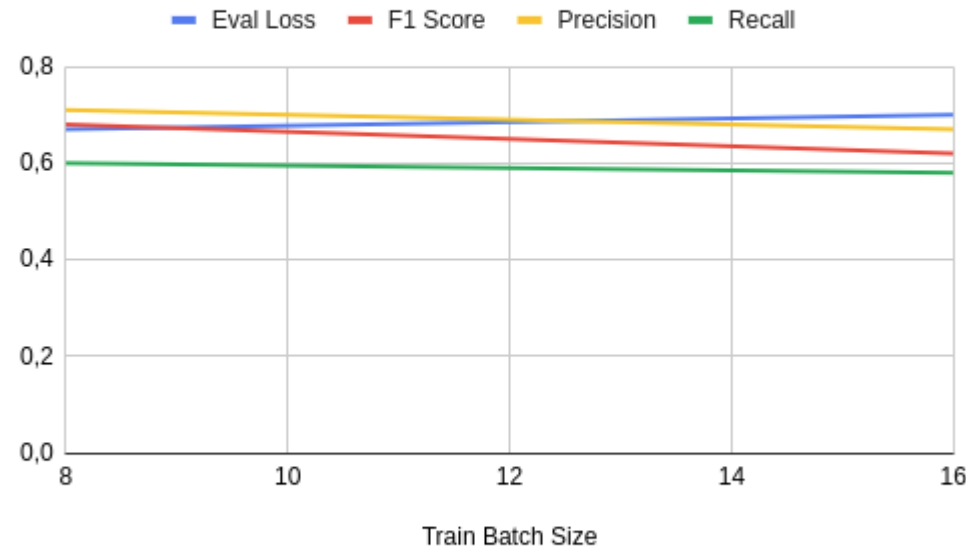
			Train Epoch	5
			Train Batch Size	8
			Eval Batch Size	32



Parametre Analizi - 3

Train Batch Size	Eval Loss	F1 Score	Precision	Recall
8	0,67	0,68	0,71	0,6
16	0,7	0,62	0,67	0,58

Eval Loss, F1 Score, Precision ve Recall

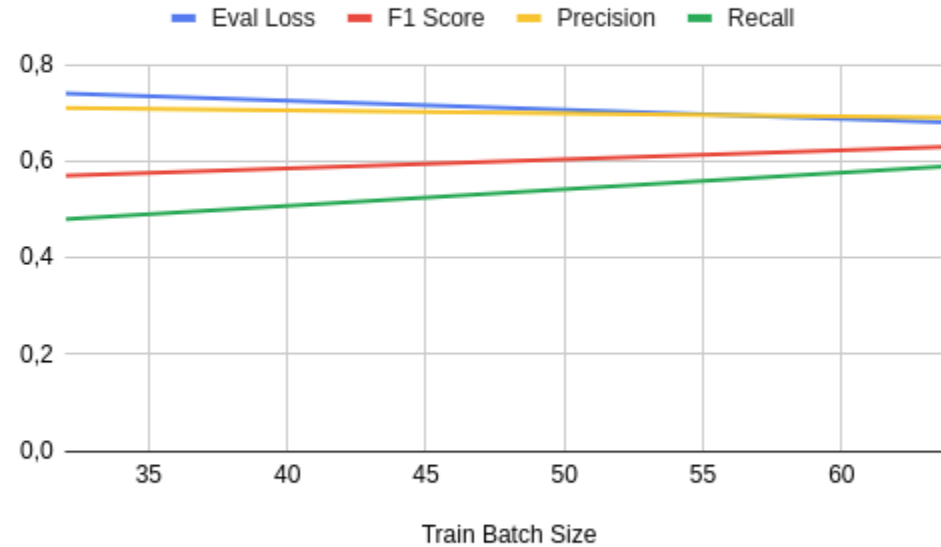


Train Epoch	5
Learning Rate	3,00E-05
Eval Batch Size	32



Eval Batch Size	Eval Loss	F1 Score	Precision	Recall
32	0,74	0,57	0,71	0,48
64	0,68	0,63	0,69	0,59

Eval Loss, F1 Score, Precision ve Recall



	Train Epoch	5
	Learning Rate	3,00E-05
	Train Batch Size	8



Tahmin Sonuçları - 1

```
[ ] prediction, model_output = model.predict(["Çevre ve Şehircilik Bakanlığı Milli Emlak Genel Müdür"])
```

prediction

100%  1/1 [00:00<00:00, 6.78it/s]

Running Prediction: 100%  1/1 [00:00<00:00, 14.97it/s]

```
[[{'Çevre': 'B-ORGANIZATION'},  
  {'ve': '0'},  
  {'Şehircilik': '0'},  
  {'Bakanlığı': 'I-ORGANIZATION'},  
  {'Milli': 'B-ORGANIZATION'},  
  {'Emlak': '0'},  
  {'Genel': 'I-ORGANIZATION'},  
  {'Müdürlüğü': 'I-ORGANIZATION'},  
  {'kadrolarında': '0'},  
  {'bulunan': '0'},  
  {'ve': '0'},  
  {'sınav': '0'},  
  {'tarihi': '0'},  
  {'itibarıyla': '0'},  
  ...]]
```



Tahmin Sonuçları - 2



```
{'ve': '0'},  
{'öğrenim': '0'},  
{'alanı': '0'},  
{'şartları': '0'},  
{'hariç': '0'},  
{'mevzuatında': '0'},  
{'öngörülen': '0'},  
{'diğer': '0'},  
{'şartları': '0'},  
{'taşıyan': '0'},  
{'personelden;': 'B-PERSON'},  
{'usul': '0'},  
{'ve': '0'},  
{'esasları': '0'},  
{'Milli': 'B-ORGANIZATION'},  
{'Emlak': '0'},  
{'Genel': 'I-ORGANIZATION'},  
{'Müdürlüğü': 'I-ORGANIZATION'},  
{'tarafından': '0'},  
{'belirlenerek': '0'},  
{'bu': '0'},  
{'maddenin': '0'},  
{'yayımlı': '0'},  
{'tarihinden': '0'},  
{'itibaren;': '0'},  
{'beş': '#'},  
{'yıl': 'B-P_TERM'},  
{'içinde': '0'},  
{'iki': '#'},  
{'defa': '0'},  
{'yapılacak': '0'},  
{'sınavda': '0'},  
{'başarılı': '0'},  
{'olanlar;': '0'},  
{'Milli': 'B-PERSON'},  
{'Emlak': '0'},  
{'Uzmanı': 'B-PERSON'},  
{'kadrolarına': '0'},  
{'atanırlar.': '0'}]]
```



1. <https://huggingface.co/dbmdz/bert-base-turkish-128k-cased>
2. <https://app.diagrams.net/>
3. <https://data.mendeley.com/datasets/cdcztymf4k/1>
4. <http://www.madeinturkeydergisi.com/kanunlar/>
5. <https://www.mevzuat.gov.tr/#kanunlar>
6. <https://www.kaggle.com/abhishek/entity-extraction-model-using-bert-pytorch>
<https://www.kaggle.com/shoumikgoswami/ner-using-random-forest-and-crf>

