



GTÜ BİL MUH BİL 495

BİL 495
İkinci İzleme

Gökhan HAS

Proje Danışmanı: Dr. Burcu YILMAZ
Aralık 2020

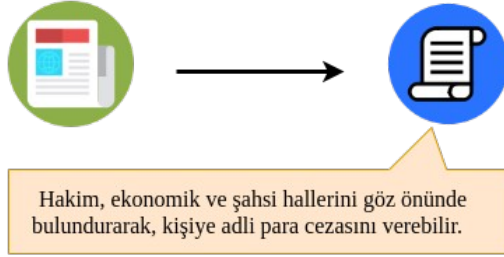


- Proje Tasarım Planı
- Veri Kümesi
- Kullanılan Modeller
- Modeller Analizi
- Tahmin Sonuçları
- Kaynaklar



Proje Tasarım Planı

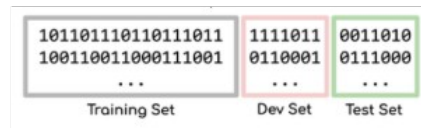
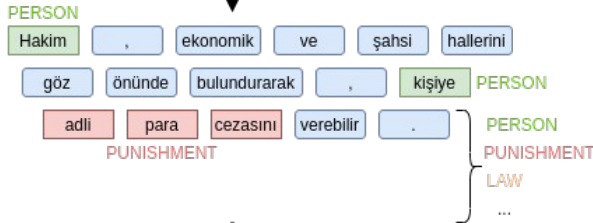
1. HAZIRLIK



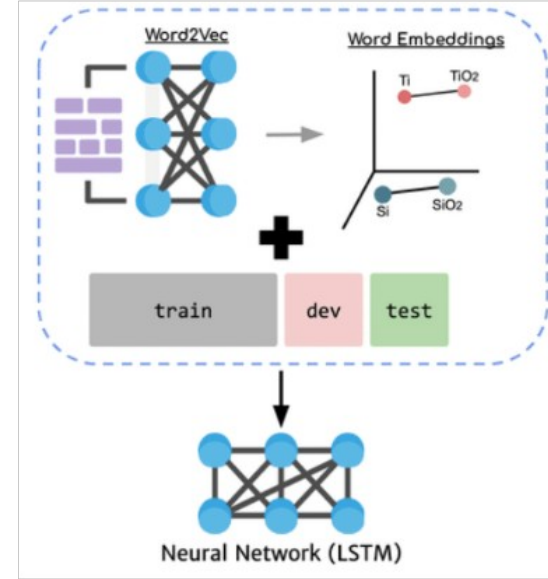
2. TOKENİZASYON



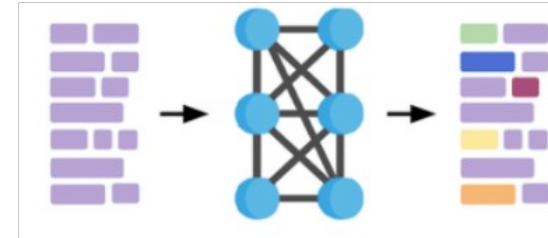
3. ETİKETLEME

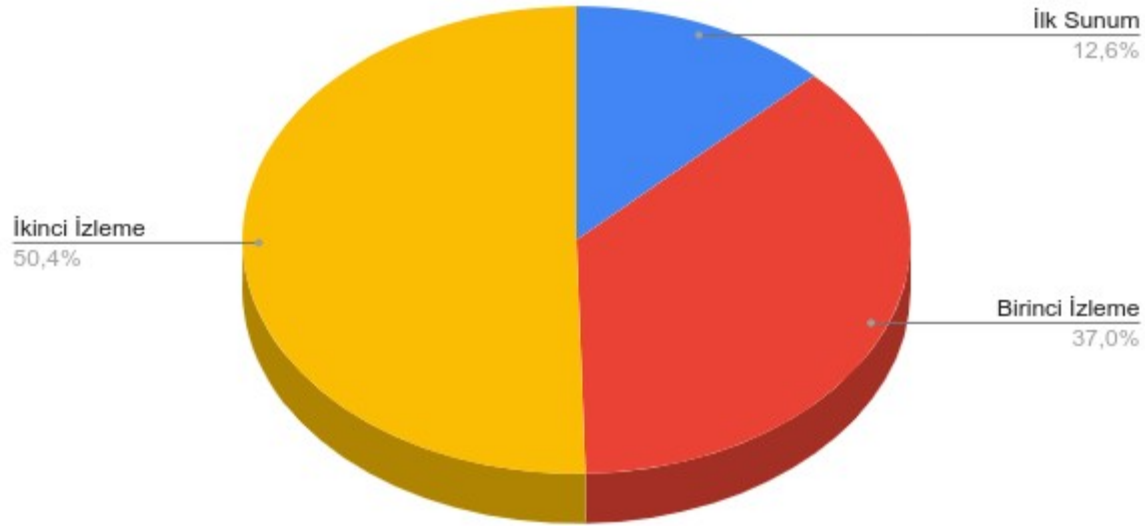


4. MODEL EĞİTİLMESİ

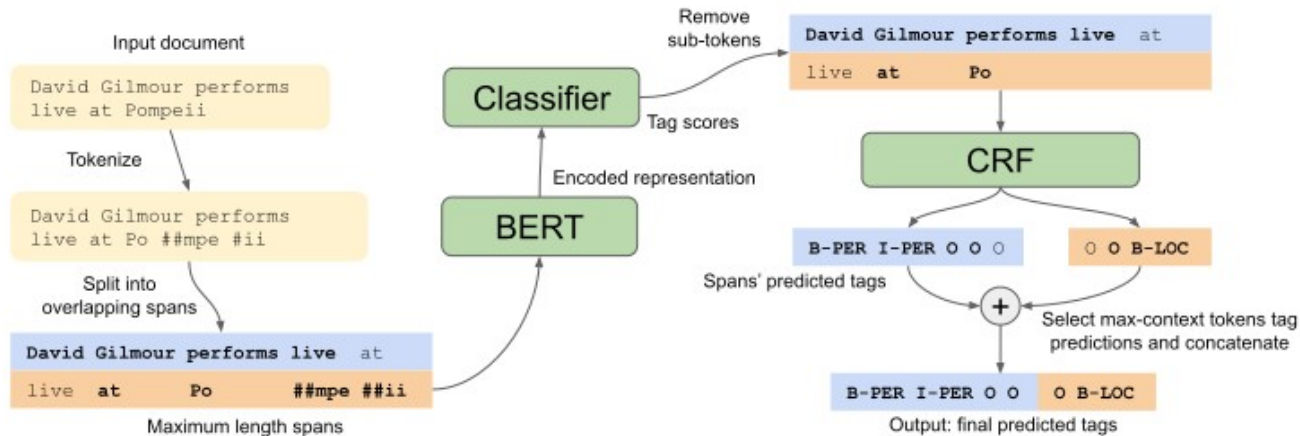


5. MODEL YARDIMIYLA SONUÇLARIN ÇIKARILMASI





Kullanılan Modeller - 1



HUGGING FACE

[Back to all models](#)

Model: `savasy/bert-base-turkish-squad`

`pytorch` `bert` `question-answering` `tr`



HUGGING FACE

[Back to all models](#)

Model: `dbmdz/bert-base-turkish-cased`

`pytorch` `tf` `bert` `tr` `license:mit`



HUGGING FACE

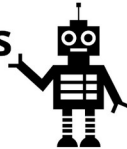
[Back to all models](#)

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

`pytorch` `tf` `bert` `tr` `license:mit`



Simple Transformers



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

MODELLERDE KULLANILAN PARAMETRELER	
Train Epoch	3
Learning Rate	3,00E-05
Train Batch Size	8
Eval Batch Size	32



KULLANILAN BERT MODELİ : bert-base-turkish-squad			
	Veri Kümesinin Büyüklüğü		
	15K	44K	60K
Eval Loss	0,65	0,47	0,33
F1 Score	0,67	0,73	0,77
Precision	0,66	0,72	0,77
Recall	0,65	0,71	0,75

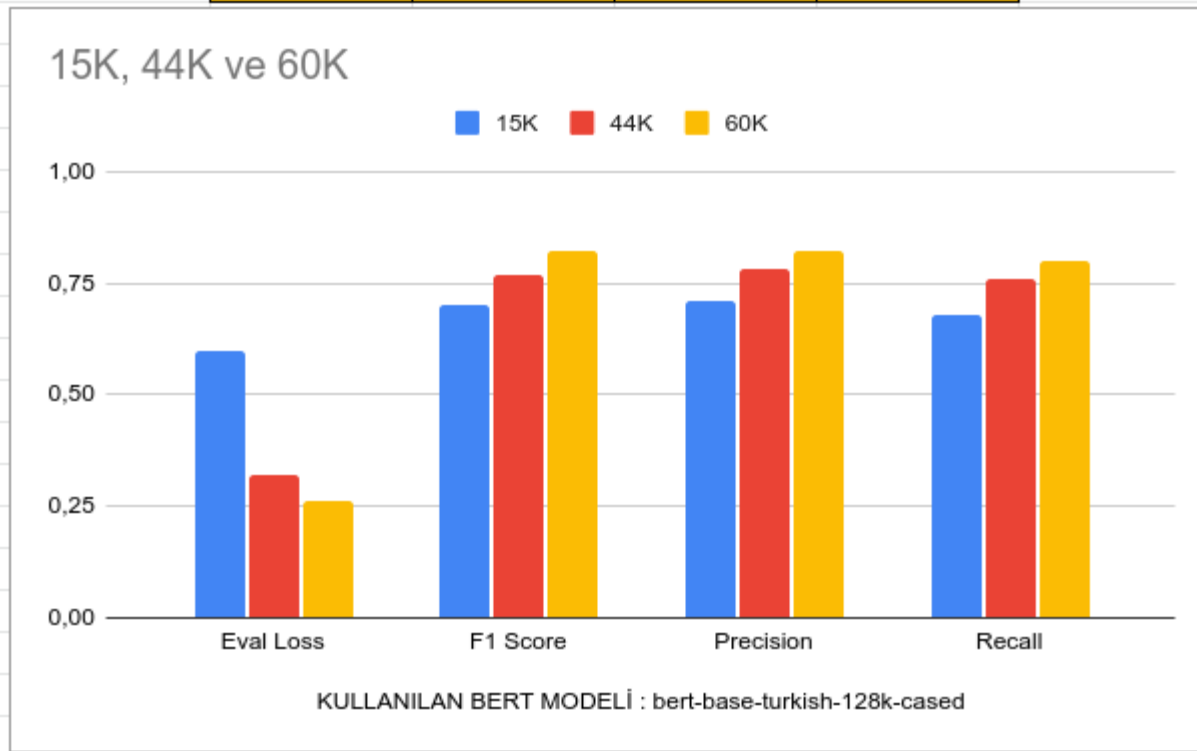


KULLANILAN BERT MODELİ : bert-base-turkish-cased			
	Veri Kümesinin Büyüklüğü		
	15K	44K	60K
Eval Loss	0,64	0,45	0,32
F1 Score	0,66	0,71	0,77
Precision	0,66	0,72	0,78
Recall	0,64	0,71	0,76



Modeller Analizi - 3

KULLANILAN BERT MODELİ : bert-base-turkish-128k-cased			
	Veri Kümesinin Büyüklüğü		
	15K	44K	60K
Eval Loss	0,6	0,32	0,26
F1 Score	0,7	0,77	0,82
Precision	0,71	0,78	0,82
Recall	0,68	0,76	0,8



Tahmin Sonuçları - 1

```
prediction, model_output = model.predict(["Çevre ve Şehircilik Bakanlığı /
```

```
prediction
```

100%  1/1 [00:00<00:00, 2.68it/s]

Running Prediction: 100%  1/1 [00:00<00:00, 10.82it/s]

```
[{'Çevre': 'B-ORGANIZATION'},  
 {'ve': 'I-ORGANIZATION'},  
 {'Şehircilik': 'I-ORGANIZATION'},  
 {'Bakanlığı': 'I-ORGANIZATION'},  
 {'Milli': 'B-ORGANIZATION'},  
 {'Emlak': 'I-ORGANIZATION'},  
 {'Genel': 'I-ORGANIZATION'},  
 {'Müdürlüğü': 'I-ORGANIZATION'},  
 {'kadrolarında': '0'},  
 {'bulunan': '0'},  
 {'ve': '0'},  
 {'sınav': '0'},  
 {'tarihi': '0'},  
 {'itibarıyla': '0'},  
 {'en': '0'},  
 {'az': '0'},  
 {'üç': '#'},  
 {'yıl': 'B-P_TERM'},  
 {'görev': '0'},  
 {'yapan': '0'},  
 {'uyarma': 'B-ELEMENT_OF_CRIME'},  
 {'ve': '0'},  
 {'kınama': 'B-ELEMENT_OF_CRIME'},  
 {'hariç': '0'},  
 {'son': '0'},  
 {'üç': '#'},  
 {'yıl': 'B-P_TERM'},  
 {'içinde': '0'},  
 {'herhangi': '0'},  
 {'bir': '0'},  
 {'disiplin': 'B-PUNISHMENT'},  
 {'cezası': 'I-PUNISHMENT'},  
 {'almayan': '0'},  
 {'Milli': 'B-ORGANIZATION'},  
 {'Emlak': '0'}]
```



Tahmin Sonuçları - 2

```
{'Emlak': '0'},
{'Uzmanı': 'I-PERSON'},
{'olabilmek': '0'},
{'için': '0'},
{'yaş': '0'},
{'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': 'I-ORGANIZATION'},
{'Genel': 'I-ORGANIZATION'},
{'Müdürlüğü': 'I-ORGANIZATION'},
{'tarafından': '0'},
{'belirlenerek': '0'},
{'bu': '0'},
{'maddenin': '0'},
{'yayımı': '0'},
{'tarihinden': '0'},
{'itibaren;': '0'},
{'beş': '#'},
{'yıl': 'B-P_TERM'},
{'içinde': '0'},
{'iki': '#'},
{'defa': '0'},
{'yapılacak': '0'},
{'sınavda': '0'},
{'başarılı': '0'},
{'olanlar;': 'B-PERSON'},
{'Milli': 'B-ORGANIZATION'},
{'Emlak': '0'},
{'Uzmanı': 'I-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>

