



# GTÜ BİL MUH BİL 495

**BİL 495**  
**İlk Sunum**

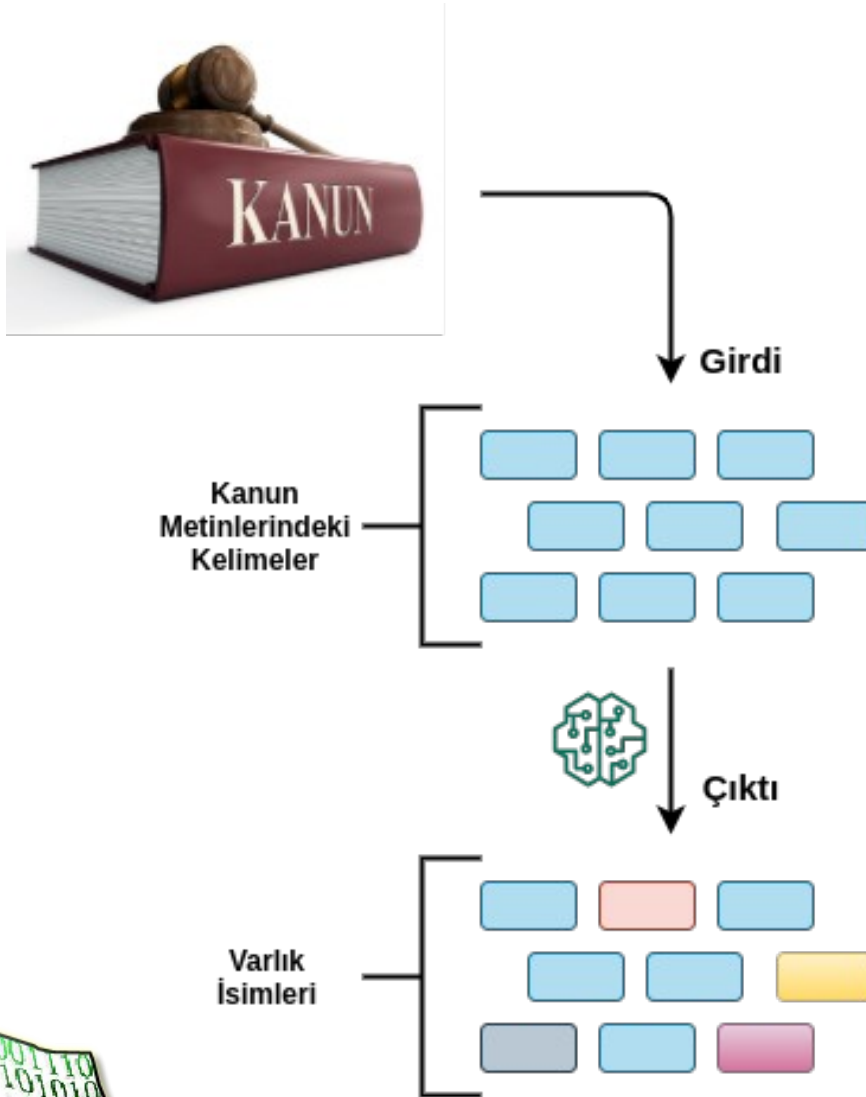
**Gökhan HAS**

**Proje Danışmanı: Dr. Burcu YILMAZ**  
**Kasım 2020**



- Projenin Şeması ve Tanımı
- Proje Tasarım Planı
- Proje Gereksinimleri
- Başarı Kriterleri
- Kaynaklar





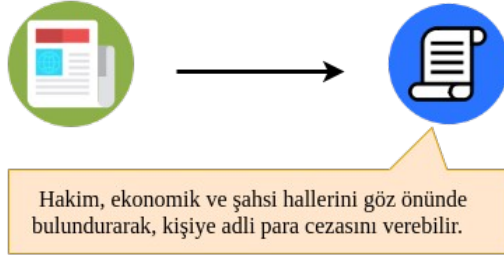
- Kanunlarda varlık isimlerinin derin öğrenme ile tespiti

Projede verilen kanun metni içeriğinden varlık isimlerinin (kişi, yer, kurum vb.) çıkarılmasının yapılması planlanmaktadır.



# Proje Tasarım Planı

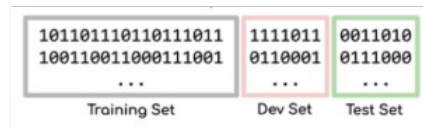
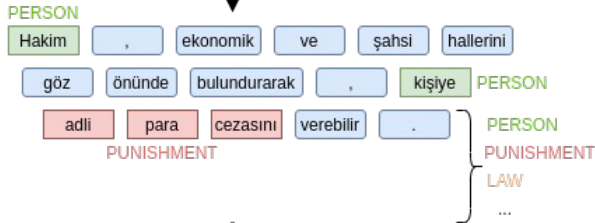
## 1. HAZIRLIK



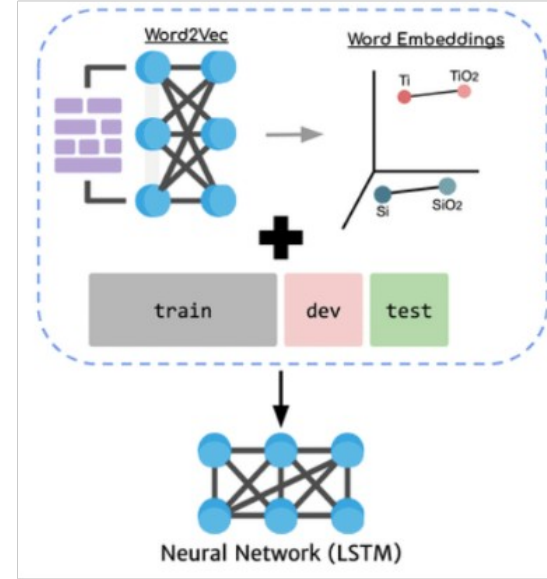
## 2. TOKENİZASYON



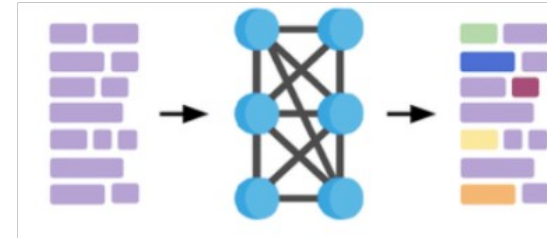
## 3. ETİKETLEME



## 4. MODEL EĞİTİLMESİ



## 5. MODEL YARDIMIYLA SONUÇLARIN ÇIKARILMASI

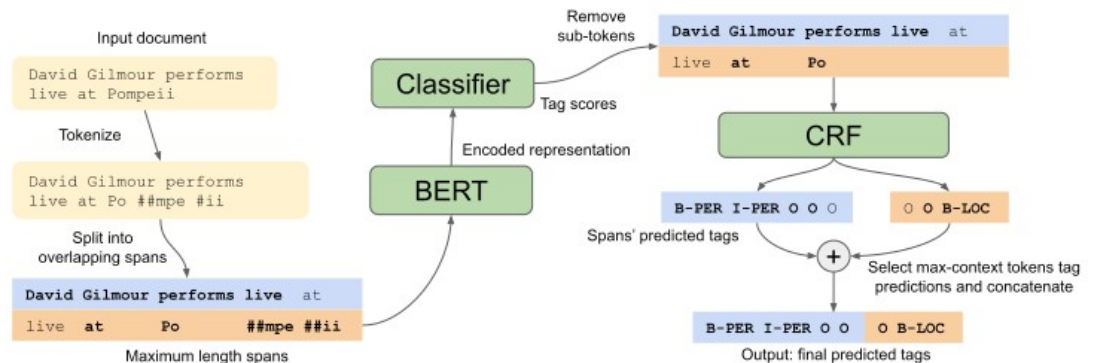
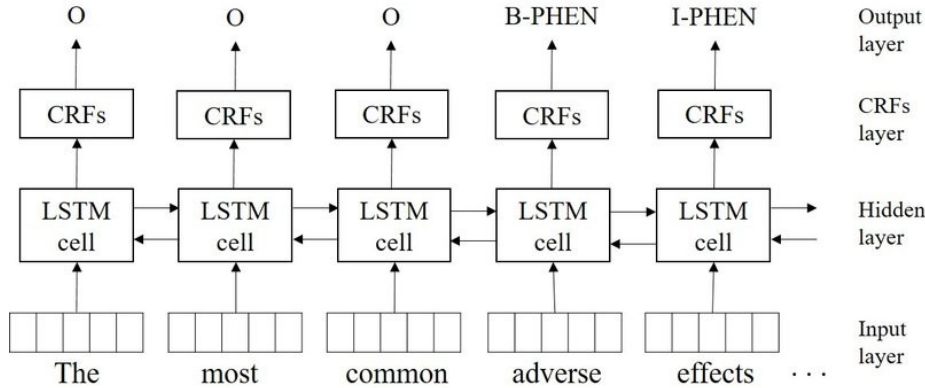


- “Kanunlarda Varlık İsimlerinin Derin Öğrenme İle Tespiti” projesi için:
  - Derin öğrenme yöntemleriyle varlık isimlerini tanıyacak algoritma geliştirmeliyim.
  - En uygun derin öğrenme algoritmasını kullanarak, projemin doğruluk oranını arttırmalıyım.
  - Kanun metinlerinden etiketlenmesi mümkün kelimeleri çıkarmalıyım.
  - Çıkardığım kelimeleri teker teker cümlelerdeki anlamlarına göre etiketleyerek bir veri seti oluşturmalıyım.
  - Oluşturduğum bu veri setinde tüm etiketlerden yeteri kadar kelime bulundurmalıyım.
  - Modelimin eğitim sonuçlarına göre veri setimi büyültüp, genişletmeliyim.



# Proje Gereksinimleri - 2

- Kullanabileceğim derin öğrenme yöntemleri
  - BERT (Bidirectional Encoder Representations from Transformers)
- - Bi-LSTM



# Proje Gereksinimleri - 3

- Yazılım dili olarak Python ve modülleri olan Tensorflow, Keras, Pytorch kütüphaneleri kullanılabilir.



TensorFlow



Keras

- Derin öğrenme modelini eğitmek ve test etmek için online olarak Google Colab veya Kaggle Kernels kullanılabilir.
- Veri kümesi olarak kanun metinlerindeki kelimelerin etiketli halleri bulunan bir veri kümesi oluşturulup, kullanılacaktır.



VALIDATION ACCURACY	F1 SCORE	VALIDATION ERROR
> 60 %	> 50 %	< 20 %

$$\frac{TP + TN}{TP + FP + TN + FN}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$F_1 = 2 * \frac{\text{precision} * \text{recall}}{\text{precision} + \text{recall}}$$





1. [https://www.kaggle.com/pendu777/bert-for-named-entity-recognition?](https://www.kaggle.com/pendu777/bert-for-named-entity-recognition?select=ner_dataset.csv)  
select=ner\_dataset.csv
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>
7. <https://www.kaggle.com/shoumikgoswami/ner-using-random-forest-and-crf>

