

Türkiye Açık Kaynak Platformu
Online Yarışma Programı

Türkçe Doğal Dil İşleme

MSKÜ-CENG-NLP-1

- EKİBİMİZ -

Alperen AYYILDIZ
Hasan ARCAS
Zehra RIDVANOĞULLARI
Murat Avzet ASAN

- EKİP ÜYELERİNİN PROJEYE SUNDUĞU KATKI -

Alperen AYYILDIZ

- Geliştirilecek modeller için gerekli literatür taraması ve dokümantasyon sağlanması.
- Modelin geliştirilmesinde Python programlama desteği.

Zehra RIDVANOĞULLARI

- Makine öğrenmesi modelleri geliştirilmesi ve model analizleri yapılması.

Murat Avzet ASAN

- Ek veri bulunması ve sınıflandırılması.

Hasan ARCAS

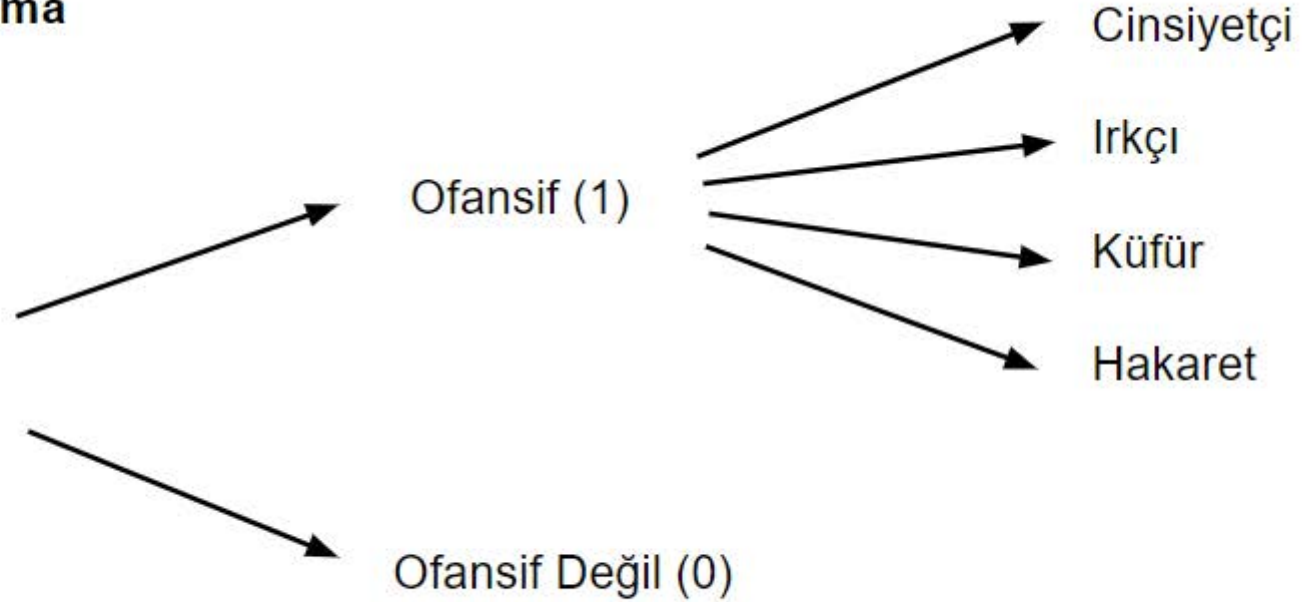
- Derin öğrenme modelleri geliştirilmesi ve model analizleri yapılması.

- PROBLEM -

- Metnin "*ofansif*" veya "*ofansif değil*" şeklinde sınıflandırması.
- *Ofansif* metinlerin alt kırılımlarına sınıflandırması.

- PROBLEMİN ÇÖZÜMÜ -

İki Aşamalı Sınıflandırma



- HANGİ YÖNTEMLE ÇÖZÜM GELİŞTİRİLDİ?-

- BERT (dbmz/bert-base-turkish-cased)
- Turkish NLP Community
- Turkish OSCAR corpus- Wikipedia- OPUS corpora

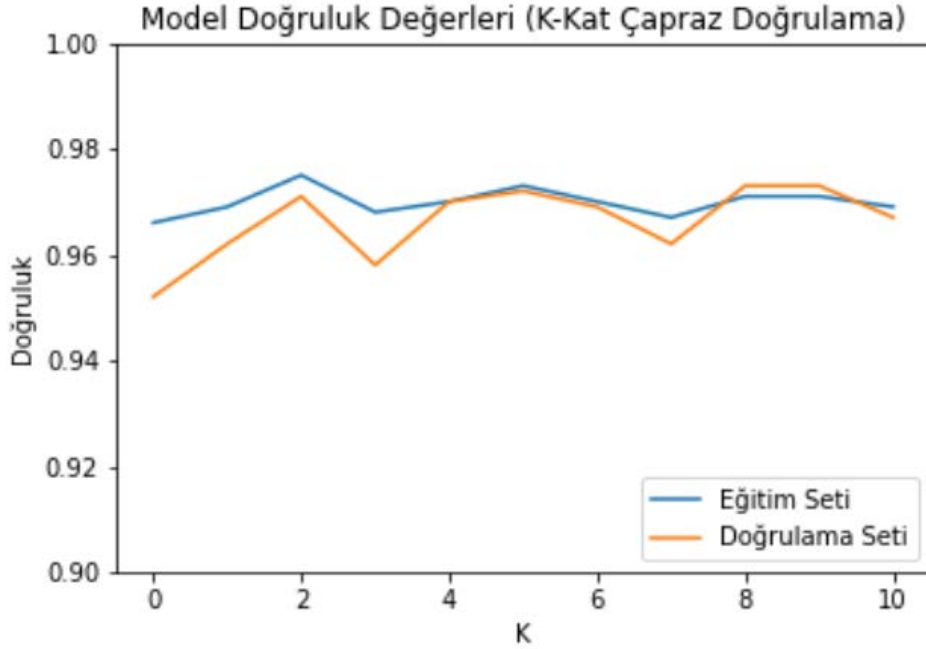
- TEKNİK ÇALIŞMALAR -

1. **Data Preprocessing (Veri Önileme)**
Lowercase, Data Cleaning
~~Stemming (Zemberek)~~
2. **Data Augmentation (Veri Arttırma) (2. model)**
3. **Model Training (Model Eğitimi)**
4. **n-Cross Validation (Çapraz Doğrulama)**

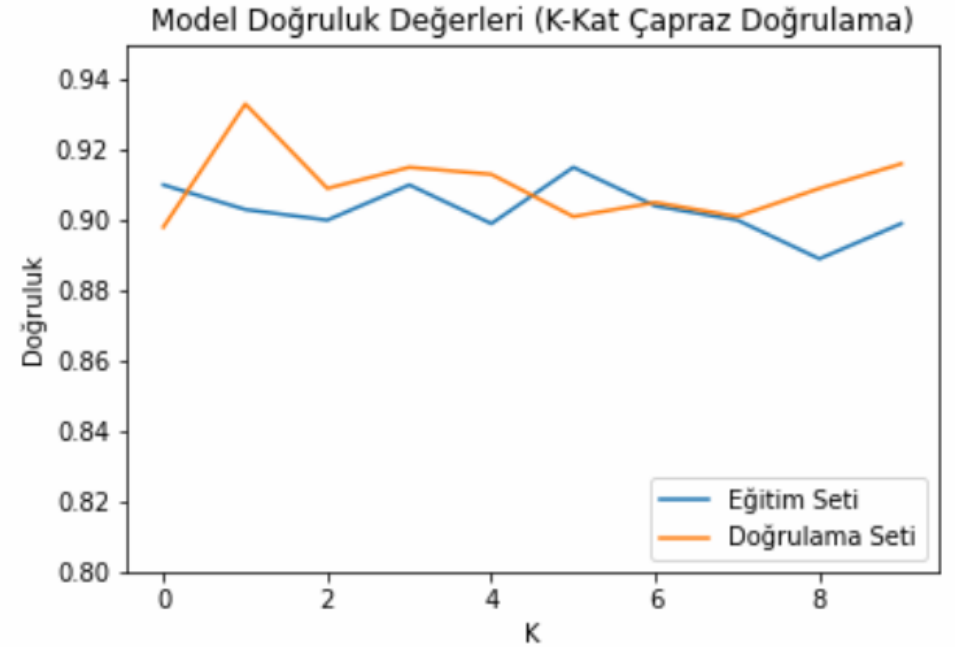
- PROJE İŞ AKIŞI -

1. Literatür Taraması
2. Data Preprocessing (Veri Önışleme)
3. Makine Öğrenmesi ve İstatiksel Modeller
4. Derin Öğrenme Modelleri
5. Dil Modeli (BERT)
6. Ek Veri Bulma/Yaratma
7. Model Training (Model Eğitimi)
8. Model Evaluation
n-Cross Validation (Çapraz Doğrulama)

SONUÇLARIMIZ

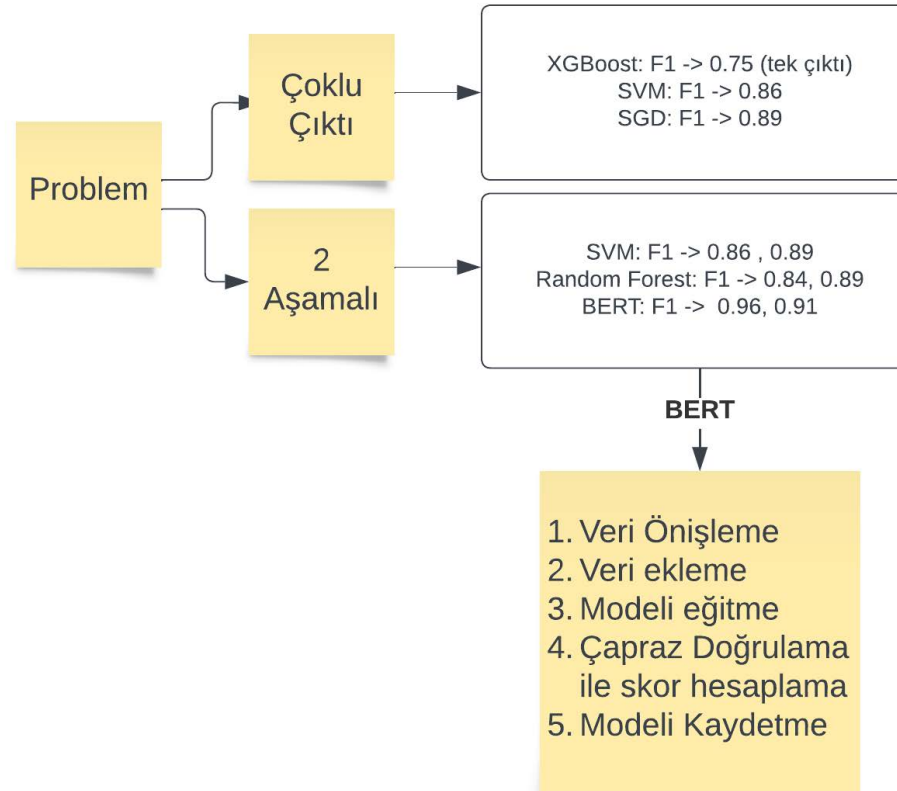


Eğitim Seti Ortalama Doğruluk Skoru ~ 0.969
Doğrulama Seti Ortalama Doğruluk Skoru ~ 0.966
Test Veri Seti F1-Skor ~ 0.972

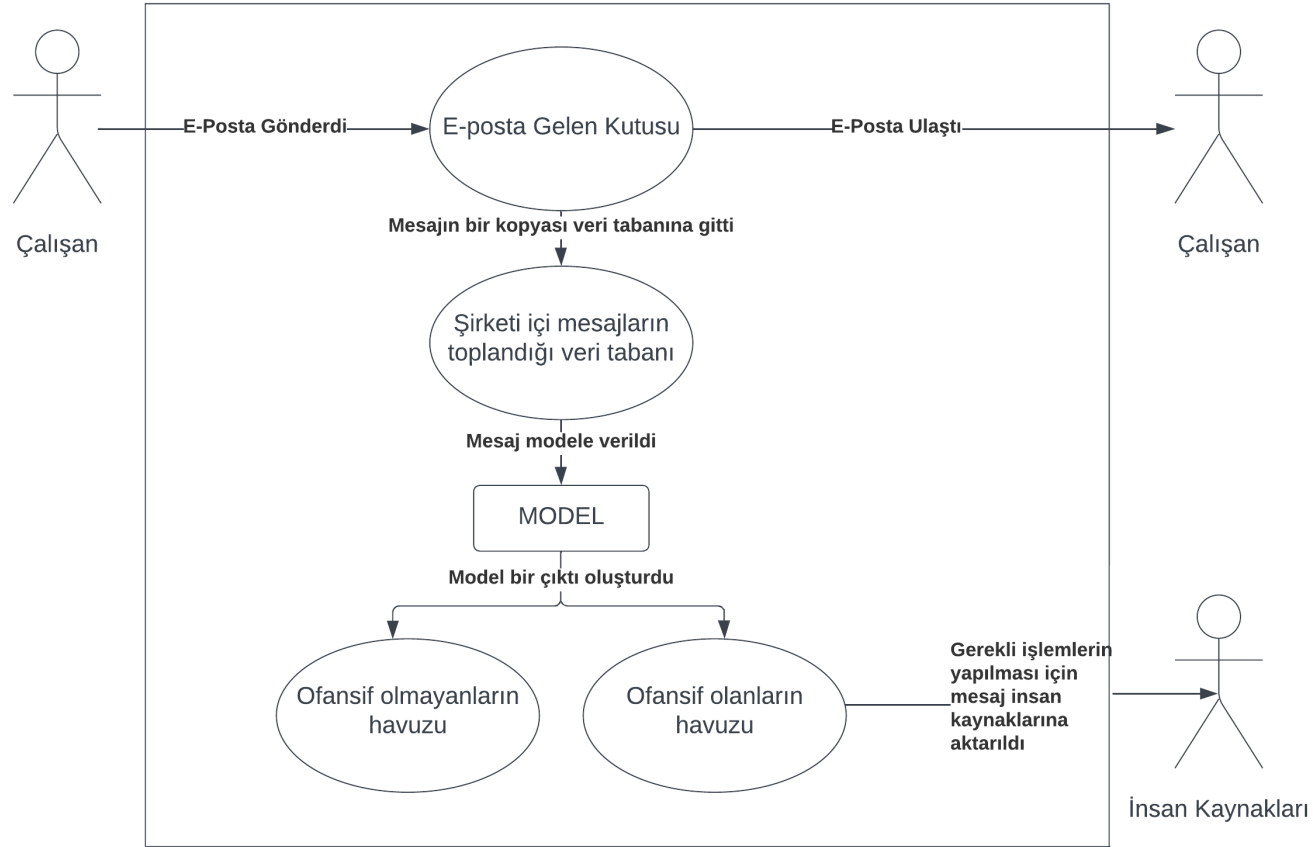


Eğitim Seti Ortalama Doğruluk Skoru ~ 0.902
Doğrulama Seti Ortalama Doğruluk Skoru ~ 0.910
Test Veri Seti F1-Skor ~ 0.905

- PROJE İLE İLGİLİ YOL HARİTAMIZ-



USE-CASE DİYAGRAMI



LINKLERİMİZ

Demo Video Linki

Github Linki

Not: Projeye ait Github Linkini ekleyiniz.

www.turkiyeacikkaynakplatformu.com



- Teşekkürler -