

# TÜRKİYE AÇIK KAYNAK PLATFORMU

ONLINE YARIŞMA PROGRAMI

Türkçe Doğal Dil İşleme

**Takım:** L2 Regülasyon

<http://www.turkiyeacikkaynakplatformu.com/>



# L2 REGÜLASYON



**Danışman**  
Ahmet Erdem



## Takım Üyeleri

Anıl Öztürk

Tarık Karakaş

İsmail Denizli



### OTTO – Multi-Objective Recommender System

Build a recommender system based on real-world e-commerce sessions  
Featured · 2587 Teams · 2 months ago

28/2587



### Foursquare - Location Matching

Match point of interest data across datasets  
Featured · Code Competition · 1079 Teams · 9 months ago

55/1079



### Anadolu Hayat Emeklilik Datathon // Coderspace

Anadolu Hayat Emeklilik Datathon // Coderspace  
Community · 114 Teams · a year ago

1/114



### EnerjiSA Üretim Hackathon

EnerjiSA Üretim Hackathon  
Community · 128 Teams · a year ago

1/128



### Migros Datathon // Coderspace

Migros Datathon // Coderspace  
Community · 186 Teams · 4 months ago

3/186



### EnerjiSA Üretim Fiziksel Hackathon

EnerjiSA Üretim Fiziksel Hackathon  
Community · 14 Teams · a year ago

3/14



# L2 REGÜLASYON

## Ekip Üyelerinin Katkısı

**Anıl Öztürk:** Eğitimlerin gerçekleşmesi ve raporlanması, kütüphane formatında bir kod tabanının oluşturulması, BERT mimarisinin implementasyonu, dökümantasyon

**Tarık Karakaş:** Ek veri taraması, literatür taraması, keşifsel veri analizi, 2 aşamalı embedding mimarisinin implementasyonu, dökümantasyon

**İsmail Denizli:** Stack-ensemble deneyleri, detaylı analizlerin çıkarımı, dökümantasyon

**Ahmet Erdem:** Domain ve mimari modelleme hakkındaki püf noktalarda teknik destek

# PROBLEM

İnternet üzerinde, kullanıcı girdilerine açık mecralarda başkalarını aşağılayan veya kötü etkileyen ofansif içerikler bulunmakta.

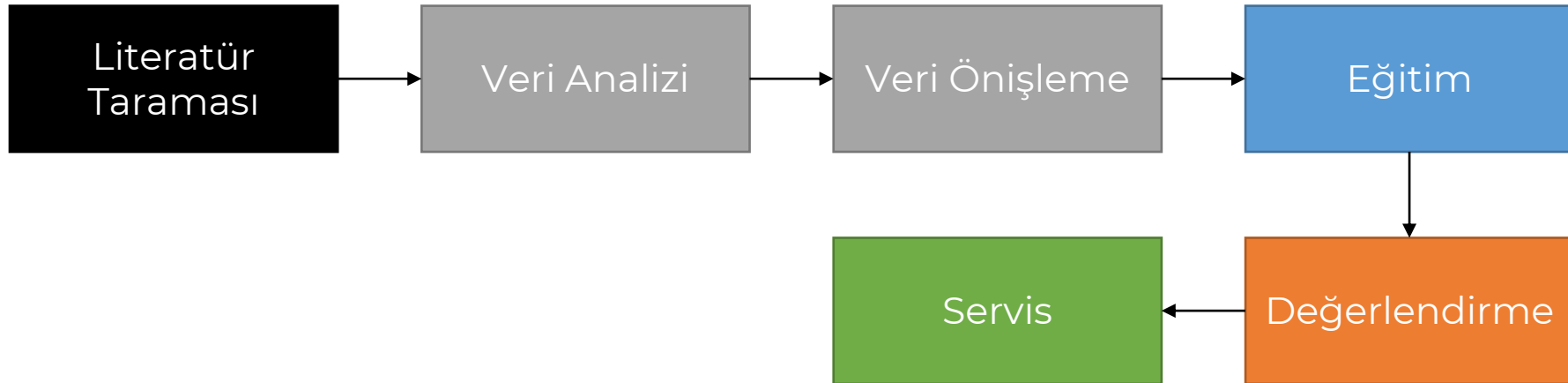
# PROBLEM

İnternet üzerinde, kullanıcı girdilerine açık mecralarda başkalarını aşağılayan veya kötü etkileyen ofansif içerikler bulunmakta.

# ÇÖZÜM

Cinsiyetçi, ırkçı, küfür ve hakaret söylemleri gibi aşağılayıcı söylemler içeren cümlelerin doğal dil işleme yöntemleri ile tespit edilmesi ve ortaya çıkan teknik yetkinliğin sektörel kullanım alanları üzerine öneriler iletilmesi.

# PROJE İŞ AKIŞI



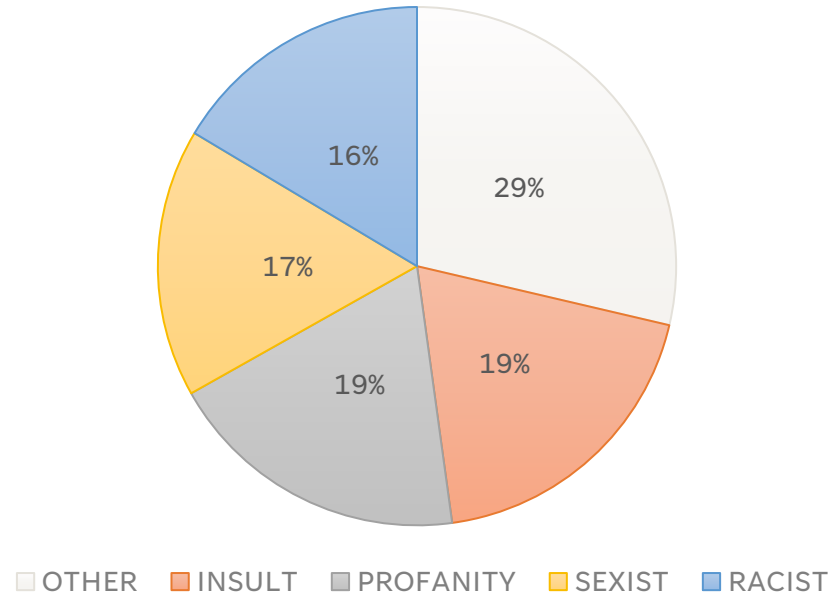
# LİTERATÜR TARAMASI

Yazarlar	Sınıflar	Mimari	Sonuç
Aken B. V., Risch J. Krestel R. ve Löser A. (2018)	Other Toxic Obscene Insult Identity Hate Severe Toxic Threat	CNN LSTM GRU Logistic Regression	En iyi sonucu Bidirectional GRU Attention (FastText) modeli ile elde etmişlerdir
Jhaveri M., Ramaiya D. ve Chadha H. S (2022)	Toxicity Obscene Sexual Explicit Identity Attack Insult Threat	BERT RNN XLNET	Focall Loss ile eğitilmiş RoBERTa modeli AUROC ve F1 olarak en iyi sonucu vermiştir
Duchêne C., Jamet H., Guillaume P. ve Dehak R. (2023)	Abusive Not Abuse	BERT ailesinden 18 farklı model	xlm-roberta-large en iyi sonucu veren model olmuştur.

<https://github.com/L2-Regulasyon/Teknofest2023/blob/main/src/models/README.md>

# VERİ ANALİZİ

Etiket Oranı

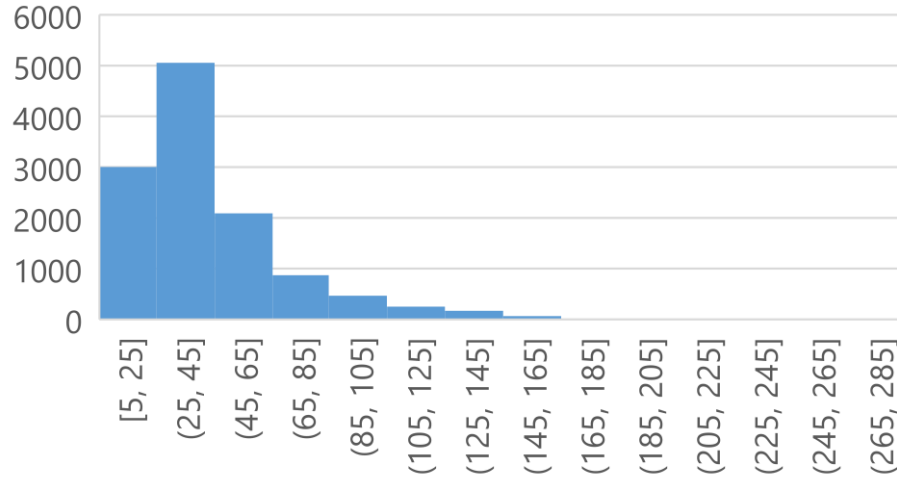


<https://github.com/L2-Regulasyon/Teknofest2023/blob/main/analysis/EDA.ipynb>

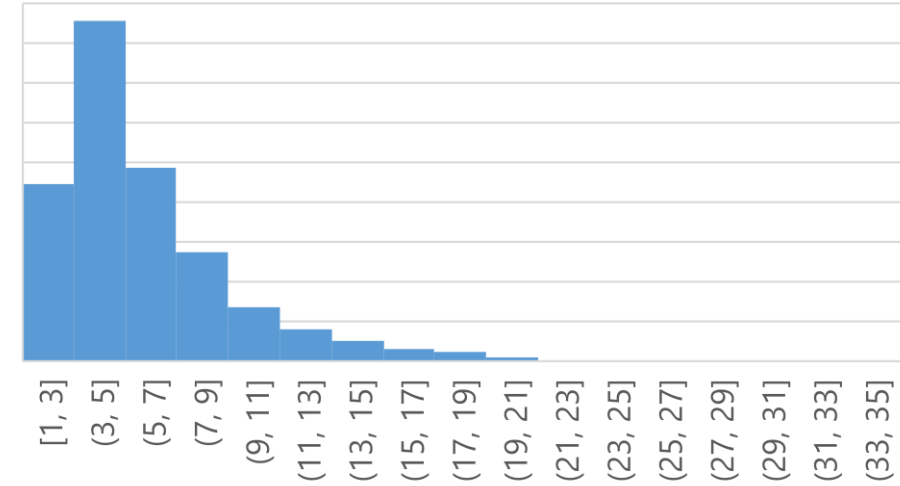


# VERİ ANALİZİ

Cümlelerdeki karakter sayısı



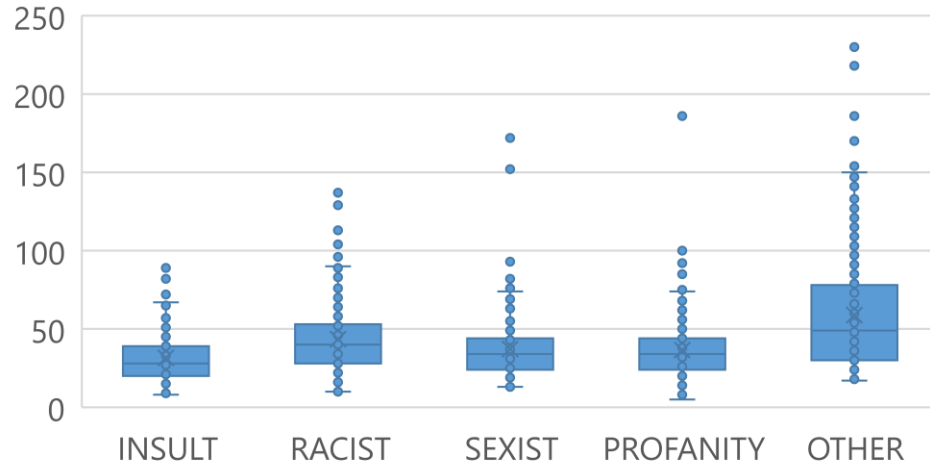
Cümlelerdeki kelime sayısı



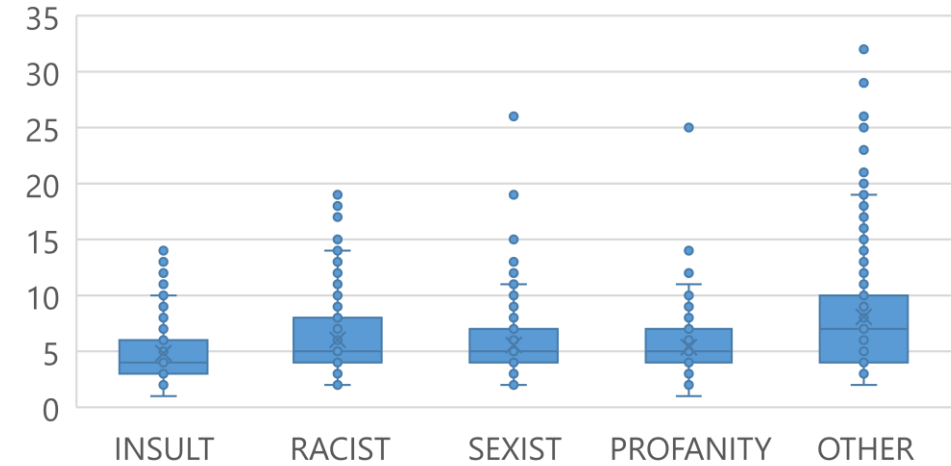
<https://github.com/L2-Regulasyon/Teknofest2023/blob/main/analysis/EDA.ipynb>

# VERİ ANALİZİ

Sınıflara göre cümlelerin karakter sayısı



Sınıflara göre cümlelerin kelime sayısı



<https://github.com/L2-Regulasyon/Teknofest2023/blob/main/analysis/EDA.ipynb>

# VERİ ANALİZİ

	OTHER	PROFANITY	RACIST	SEXIST	INSULT
gibi	0,003	0,006	0,016	0,019	0,021
sen	0,000	0,004	0,003	0,005	0,019
ne	0,003	0,004	0,007	0,008	0,014
kadar	0,003	0,006	0,003	0,003	0,009
o	0,005	0,005	0,003	0,003	0,009
çok	0,006	0,002	0,005	0,003	0,007
bile	0,001	0,002	0,002	0,001	0,007
sana	0,000	0,004	0,001	0,001	0,007
senden	0,000	0,001	0,000	0,001	0,006
insan	0,001	0,001	0,001	0,000	0,005

	SEXIST	INSULT	PROFANITY	OTHER	RACIST
suriyeliler	0,000	0,000	0,000	0,000	0,009
çingene	0,000	0,000	0,000	0,000	0,008
gavur	0,000	0,000	0,000	0,000	0,007
kürt	0,000	0,000	0,000	0,000	0,006
türk	0,000	0,000	0,000	0,001	0,005
arap	0,000	0,000	0,000	0,000	0,005
ermenii	0,000	0,000	0,000	0,000	0,004
suriyeli	0,000	0,000	0,000	0,000	0,003
araplar	0,000	0,000	0,000	0,000	0,003
zenciler	0,000	0,000	0,000	0,000	0,003

	OTHER	RACIST	SEXIST	INSULT	PROFANITY
seni	0,000	0,001	0,001	0,013	0,019
senin	0,000	0,001	0,001	0,013	0,018
or**pu	0,000	0,000	0,000	0,000	0,010
ananın	0,000	0,000	0,000	0,000	0,010
ben	0,001	0,001	0,001	0,003	0,006
çocuğu	0,000	0,000	0,002	0,001	0,005
s*k	0,000	0,000	0,000	0,000	0,004
g*t	0,000	0,000	0,000	0,000	0,003
sokarım	0,000	0,000	0,000	0,000	0,003
pe***enk	0,000	0,000	0,000	0,000	0,002

	PROFANITY	OTHER	RACIST	INSULT	SEXIST
kadın	0,000	0,001	0,000	0,001	0,045
erkek	0,000	0,001	0,000	0,000	0,028
kadınlar	0,000	0,000	0,001	0,000	0,016
kız	0,000	0,001	0,001	0,000	0,014
dediğin	0,000	0,000	0,000	0,000	0,013
kadının	0,000	0,000	0,000	0,000	0,009
adam	0,001	0,001	0,001	0,003	0,009
karı	0,001	0,000	0,000	0,000	0,008
erkekler	0,000	0,000	0,000	0,000	0,008
olur	0,002	0,001	0,002	0,001	0,008

# VERİ ANALİZİ

	PROFANITY	SEXIST	INSULT	RACIST	OTHER
bir	0,011	0,014	0,019	0,020	0,028
ve	0,001	0,003	0,002	0,010	0,026
bu	0,004	0,006	0,013	0,011	0,018
için	0,001	0,003	0,003	0,003	0,008
da	0,003	0,003	0,003	0,004	0,007
de	0,004	0,004	0,004	0,004	0,006
ile	0,001	0,001	0,001	0,002	0,005
olarak	0,000	0,001	0,000	0,002	0,004
her	0,001	0,004	0,002	0,003	0,004
ilgili	0,000	0,000	0,000	0,000	0,002

# VERİ ANALİZİ

## Sınıf – Alt Sınıf uyumu

	is_offensive		
	Sayı	Ortalama	Toplam
INSULT	2419	0.995	2407
OTHER	3616	0.019	72
PROFANITY	2398	0.994	2386
RACIST	2072	0.991	2054
SEXIST	2112	0.993	2099

<https://github.com/L2-Regulasyon/Teknofest2023/blob/main/analysis/EDA.ipynb>

# VERİ ANALİZİ

## Büyük harf kullanımı

Sınıf	Ortalama büyük harf içermeye oranı
HEPSİ	% 30.26
OTHER	% 0.05
INSULT	% 38.06
PROFANITY	% 26.94
RACIST	% 42.45
SEXIST	% 49.64

<https://github.com/L2-Regulasyon/Teknofest2023/blob/main/analysis/EDA.ipynb>

# VERİ ANALİZİ

## Benzer cümlelerin tespiti

**Metot:** LCS (Longest Common Subsequence)

Cümle 1	Cümle 2
Çingene ile ahbap olursan mükafatın kalbur olur	Çingeneyle ahbap olanın mükafatı kalbur olur
Senin bana bu yaptığını Yunan gavuru yapmazdı	senin bu yaptığını yunan gavuru yapmaz
seni param ile satın alırım	Seni paramla satın alırım
kadın dediğin cilveli olacak	Kadın dediğin cilveli olur

~1000 benzer cümle çifti

<https://github.com/L2-Regulasyon/Teknofest2023/blob/main/analysis/Duplicate%20Analysis.ipynb>

# VERİ ÖNİŞLEME

## Denenen Metotlar

- **Unicode normalization:** Karakterlerin Unicode formatına göre düzeltilmesi
- **Zemberek normalization:** Hatalı olabilecek yazımların düzeltilmesi
- **Punctuation removal:** Noktalama işaretlerinin silinmesi
- **Küçük karakter çevrimi:** Karakterlerin küçültülmesi
- **Büyük harf işaretçisi:** Büyük harflerin başlarına özel işaretçi (#) eklenmesi
- **Cümle uzunluk işaretçisi:** Cümle uzunluğunun başa eklenen özel işaretçilerle belirtilmesi

[https://github.com/L2-Regulasyon/Teknofest2023/blob/main/src/utils/preprocess\\_utils.py](https://github.com/L2-Regulasyon/Teknofest2023/blob/main/src/utils/preprocess_utils.py)



# VERİ ÖNİŞLEME

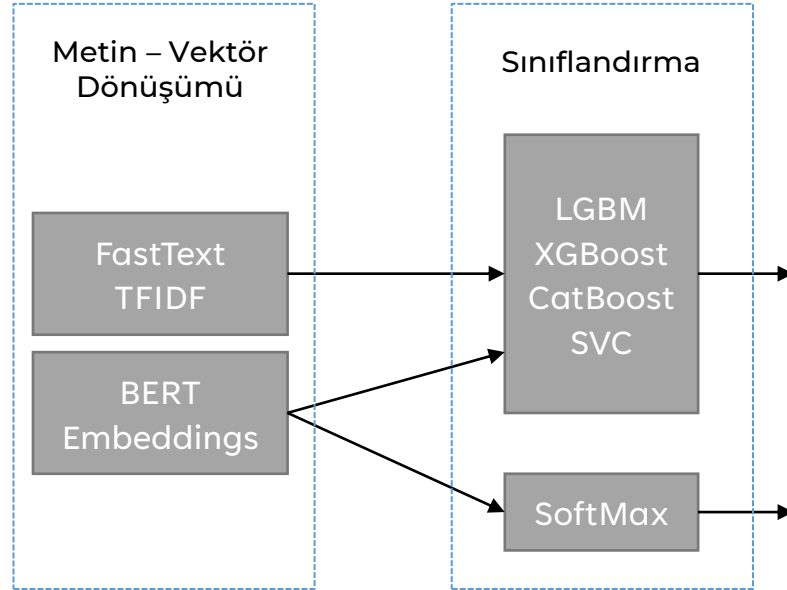
## Denenen Metotlar

- **Unicode normalization:** Karakterlerin Unicode formatına göre düzeltilmesi
- **Zemberek normalization:** Hatalı olabilecek yazımların düzeltilmesi
- **Punctuation removal:** Noktalama işaretlerinin silinmesi
- **Küçük karakter çevrimi:** Karakterlerin küçültülmesi
- **Büyük harf işaretçisi:** Büyük harflerin başlarına özel işaretçi (#) eklenmesi
- **Cümle uzunluk işaretçisi:** Cümle uzunluğunun başa eklenen özel işaretçilerle belirtilmesi

[https://github.com/L2-Regulasyon/Teknofest2023/blob/main/src/utls/preprocess\\_utils.py](https://github.com/L2-Regulasyon/Teknofest2023/blob/main/src/utls/preprocess_utils.py)

# MODEL EĞİTİMİ

## Denenen Mimariler

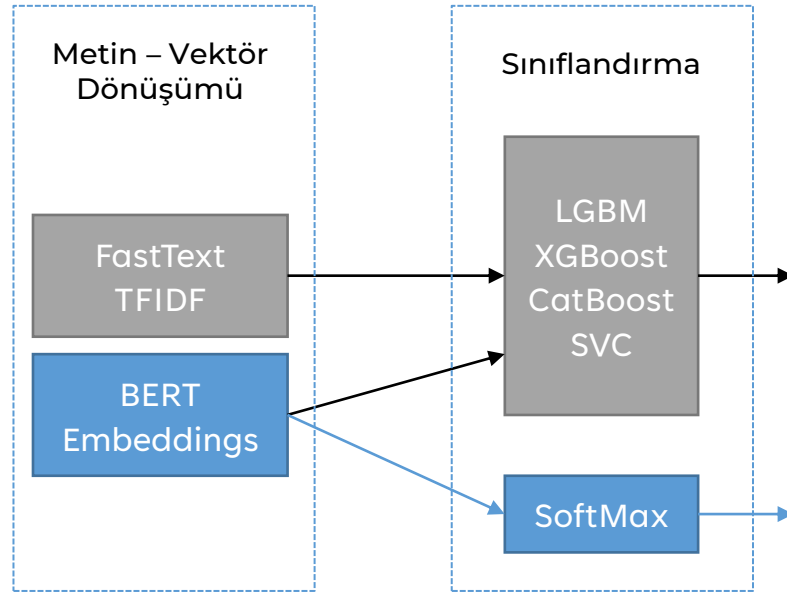


Model	F1-Macro	F1-OTHER	F1-INSULT	F1-RACIST	F1-SEXIST	F1-PROFANITY
toxic-dbmdbz-bert-base-turkish-128k-uncased	95.58	96.63	92.16	96.67	96.43	95.99
dbmdz-bert-base-turkish-128k-uncased (Fine-Tuned) Embeddings + svc	95.54	96.59	92.14	96.71	96.28	95.98
dbmdz-bert-base-turkish-128k-uncased (Fine-Tuned) Embeddings + lgbm	95.5	96.62	91.94	96.6	96.34	96.01
dbmdz-bert-base-turkish-128k-uncased (Fine-Tuned) Embeddings + xgb	95.48	96.59	91.94	96.52	96.41	95.95
dbmdz-bert-base-turkish-128k-uncased (Fine-Tuned) Embeddings + catboost	95.44	96.51	91.91	96.69	96.15	95.95
xlm-roberta-base (Fine-Tuned) Embeddings + lgbm	92.92	94.35	87.27	94.37	94.66	93.96
xlm-roberta-base (Fine-Tuned) Embeddings + svc	92.89	94.24	87.43	94.31	94.48	93.97
xlm-roberta-base (Fine-Tuned) Embeddings + xgb	92.84	94.29	87.21	94.38	94.46	93.87

<https://github.com/L2-Regulasyon/Teknofest2023/tree/main/src/models>

# MODEL EĞİTİMİ

## Denenen Mimariler



Model	F1-Macro	F1-OTHER	F1-INSULT	F1-RACIST	F1-SEXIST	F1-PROFANITY
toxic-dbmdbz-bert-base-turkish-128k-uncased	95.58	96.63	92.16	96.67	96.43	95.99
dbmdz-bert-base-turkish-128k-uncased (Fine-Tuned) Embeddings + svc	95.54	96.59	92.14	96.71	96.28	95.98
dbmdz-bert-base-turkish-128k-uncased (Fine-Tuned) Embeddings + lgbm	95.5	96.62	91.94	96.6	96.34	96.01
dbmdz-bert-base-turkish-128k-uncased (Fine-Tuned) Embeddings + xgb	95.48	96.59	91.94	96.52	96.41	95.95
dbmdz-bert-base-turkish-128k-uncased (Fine-Tuned) Embeddings + catboost	95.44	96.51	91.91	96.69	96.15	95.95
xlm-roberta-base (Fine-Tuned) Embeddings + lgbm	92.92	94.35	87.27	94.37	94.66	93.96
xlm-roberta-base (Fine-Tuned) Embeddings + svc	92.89	94.24	87.43	94.31	94.48	93.97
xlm-roberta-base (Fine-Tuned) Embeddings + xgb	92.84	94.29	87.21	94.38	94.46	93.87

<https://github.com/L2-Regulasyon/Teknofest2023/tree/main/src/models>

# MODEL EĞİTİMİ

## Seçilen Mimari



**dbmdz/bert-base-turkish-128k-uncased**

<https://huggingface.co/dbmdz/bert-base-turkish-128k-uncased>

- **Eğitim Verisi:** 45 milyar token (~35GB)
- **Sözlük Boyutu:** 128000 token

# MODEL EĞİTİMİ

## Eğitim Ayarları

- **Stratified Batch Sampling:** Örneklenen mini-batchlerde sınıf dengesini sağlar
- **Early-Stop Kapalı Eğitim:** Doğrulama setine dolaylı yoldan overfit olunması engellenir
- **Online Hard Example Mining:** Her batch'te sürekli en zor yüzdelik dilime ait örneklerin loss'u değerlendirilir
- **Sınıf Ağırlıklandırma:** Sınıfların loss çarpanı örnek sayılarına göre ayarlanır
- **Cosine LR Scheduler:** Fine-tuning için azdan başlayıp azda biten bir LR zamanlayıcı kullanılır
- **Gradient Clipping:** Gradyan patlaması yaşamamak adına hesaplanan gradyanlar kırpılır
- **Weight Decay:** Loss fonksiyonuna regülerizasyon eklenir
- **Label Smoothing:** Modelin aşırı emin tahminler üretmesini engellemek adına gerçek etiket değerleri bulanıklaştırılır
- **LLRD:** Modelin encoder yapısında baştan sonra giderken azalan learning-rate uygulanır
- **MLM Pre-Training:** Modelin görev metinlerine alışması için maskeleme kullanan bir ön eğitim uygulanır

# MODEL EĞİTİMİ

## Kullanılmayan Geliştirmeler

- **Voting Ensemble:** Her CV fold'unda eğitilen modelin ayrı olarak kaydedilip sınıf olasılık tahminlerinin ortalamasının alınması

[https://github.com/L2-Regulasyon/Teknofest2023/blob/main/src/models/voting\\_bert\\_model.py](https://github.com/L2-Regulasyon/Teknofest2023/blob/main/src/models/voting_bert_model.py)

- **OOF Stacking:** Modellerin OOF tahminleri üzerinden olası hatalı tahminleri öğrenip düzelten GBDT mimarili modellerin kullanılması
- **Sentiment Analysis Feature:** Verilen metnin pozitif-negatif içerik skorunun ön eğitilmiş bir modelden alınarak model girdisi olarak kullanılması

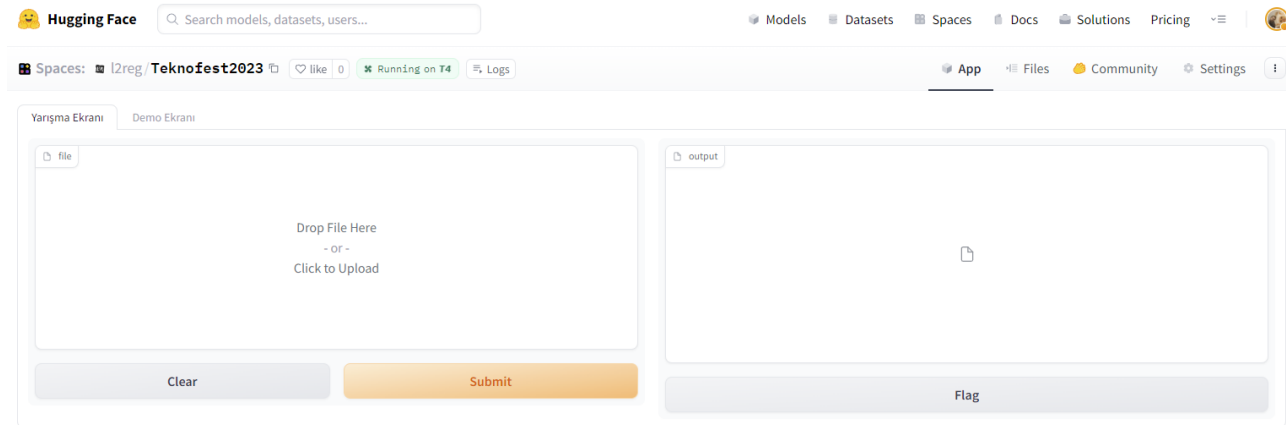
[https://github.com/L2-Regulasyon/Teknofest2023/blob/main/temp\\_notebooks/Anil%20-%20Stage2%20Modeling.ipynb](https://github.com/L2-Regulasyon/Teknofest2023/blob/main/temp_notebooks/Anil%20-%20Stage2%20Modeling.ipynb)

# MODEL DEĞERLENDİRME

- **Stratified Cross-Validation:** Sınıfların bütün eğitim ve test foldlarına eşit oranda dağılması sağlandı
- **Out-of-Fold Değerlendirme:** Bütün test foldlarına ait tahminler saklanarak birleştirildi. Bütün veriye ait bir tahmin seti elde edildi.
- **Model Zoo:** Yapılan bütün cross-validation içeren eğitimler skorları ve çıktıları ile saklandı. Bütün modeller eğitimler bittikten sonra rahatlıkla kıyaslanabildi.

# MODELİN SERVİSİ

- **Gradio** kullanılarak yarışma arayüzü oluşturuldu ve demo sekmesi eklendi.
- **HuggingFace** kullanılarak oluşturulan server ve eğitilen modeller herkese açık hale getirildi. Server üzerinde **NVIDIA T4** ekran kartı kullanıldı. (315 cümle / saniye)

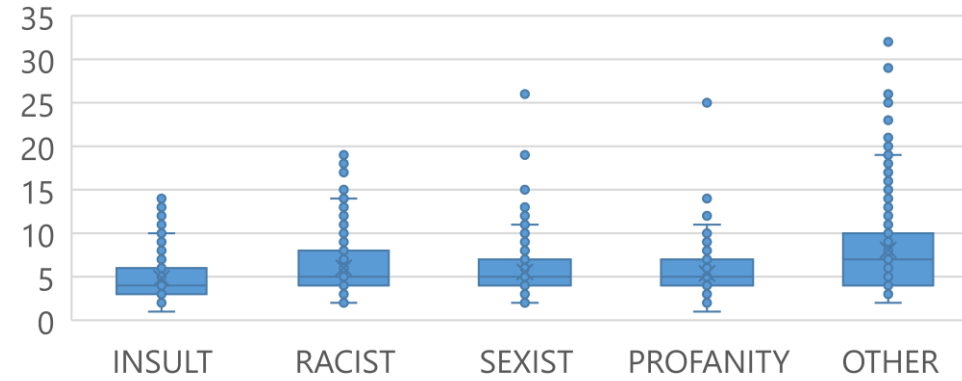


<https://huggingface.co/l2reg>



# VERİ VE MODELDEKİ ÖNYARGI

Sınıflara göre cümlelerin kelime sayısı



Sınıf	Ortalama büyük harf içerme oranı
HEPSİ	% 30.26
OTHER	% 0.05
INSULT	% 38.06
PROFANITY	% 26.94
RACIST	% 42.45
SEXIST	% 49.64

# VERİ VE MODELDEKİ ÖNYARGI

	OTHER	PROFANITY	RACIST	SEXIST	INSULT
gibi	0,003	0,006	0,016	0,019	0,021
sen	0,000	0,004	0,003	0,005	0,019
ne	0,003	0,004	0,007	0,008	0,014
kadar	0,003	0,006	0,003	0,003	0,009
o	0,005	0,005	0,003	0,003	0,009
çok	0,006	0,002	0,005	0,003	0,007
bile	0,001	0,002	0,002	0,001	0,007
sana	0,000	0,004	0,001	0,001	0,007
senden	0,000	0,001	0,000	0,001	0,006
insan	0,001	0,001	0,001	0,000	0,005

	PROFANITY	OTHER	RACIST	INSULT	SEXIST
kadın	0,000	0,001	0,000	0,001	0,045
erkek	0,000	0,001	0,000	0,000	0,028
kadınlar	0,000	0,000	0,001	0,000	0,016
kız	0,000	0,001	0,001	0,000	0,014
dediğin	0,000	0,000	0,000	0,000	0,013
kadının	0,000	0,000	0,000	0,000	0,009
adam	0,001	0,001	0,001	0,003	0,009
karı	0,001	0,000	0,000	0,000	0,008
erkekler	0,000	0,000	0,000	0,000	0,008
olur	0,002	0,001	0,002	0,001	0,008

	PROFANITY	SEXIST	INSULT	RACIST	OTHER
bir	0,011	0,014	0,019	0,020	0,028
ve	0,001	0,003	0,002	0,010	0,026
bu	0,004	0,006	0,013	0,011	0,018
için	0,001	0,003	0,003	0,003	0,008
da	0,003	0,003	0,003	0,004	0,007
de	0,004	0,004	0,004	0,004	0,006
ile	0,001	0,001	0,001	0,002	0,005
olarak	0,000	0,001	0,000	0,002	0,004
her	0,001	0,004	0,002	0,003	0,004
ilgili	0,000	0,000	0,000	0,000	0,002

# VERİ VE MODELDEKİ ÖNYARGI

<https://github.com/L2-Regulasyon/Teknofest2023/blob/main/src/README.md#112-model-data-unbiasing---kullan%C4%B1m-1---kullan%C4%B1m-2---kullan%C4%B1m-3>

Model:

- Cümleler uzadıkça **OTHER** olarak sınıflandırmaya meylediyor
- Büyük harf içeren cümleleri **ofansif** olarak sınıflandırmaya meylediyor
- Cinsiyet kelimelerini **SEXISM** ile fazla ilişkilendiriyor
- Hitabet kelimelerini **INSULT** ile eşleştiriyor

# VERİ VE MODELDEKİ ÖNYARGI ÖNERİLEN ÇÖZÜM

<https://github.com/L2-Regulasyon/Teknofest2023/blob/main/SOLUTION.md#23-final-modeller>

## **Cümleler uzadıkça OTHER olarak sınıflandırmaya meylediyor**

Eğitim sırasında OTHER sınıfına ait cümleler OTHER sınıfına ait olmayan rasgele cümlelerin uzunluklarına kırıldı.

## **Büyük harf içeren cümleleri ofansif olarak sınıflandırmaya meylediyor**

Büyük harflere özel işaretçi kullanılmadan küçültme yapılabilir.

## **Cinsiyet kelimelerini SEXISM ile fazla ilişkilendiriyor**

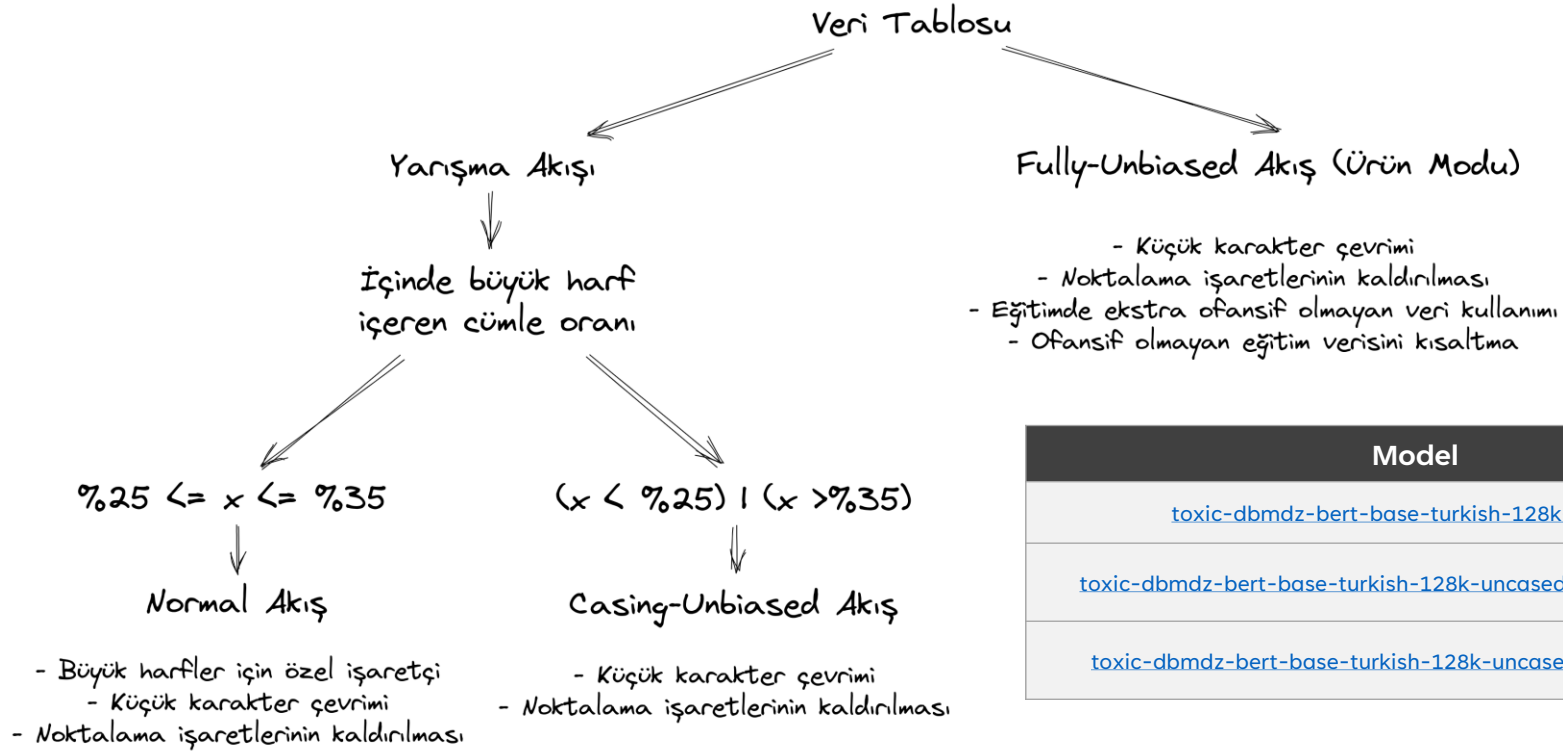
Kadın ve erkek terimleriyle ilgili ofansif olmayan metinler OTHER grubu olarak veriye eklenebilir.

## **Hitabet kelimelerini INSULT ile eşleştiriyor**

Ofansif olmayan conversational metinler OTHER sınıfı altında veriye eklenebilir.

# VERİ VE MODELDEKİ ÖNYARGI OPTİMUM İŞLEM AKIŞI

<https://github.com/L2-Regulasyon/Teknofest2023/blob/main/SOLUTION.md#23-final-modeller>



Model	F1-Macro
<a href="#">toxic-dbmddz-bert-base-turkish-128k-uncased</a>	95.58
<a href="#">toxic-dbmddz-bert-base-turkish-128k-uncased-casing-unbiased</a>	94.86
<a href="#">toxic-dbmddz-bert-base-turkish-128k-uncased-fully-unbiased</a>	93.36

# DEMO

Spaces: l2reg/ Teknofest2023 like 0 Running on T4 Logs

App Files Community Settings

Yarışma Ekranı Demo Ekranı

selected\_model

☐ Yarışma Modeli ☐ Case-Unbiased Model ☐ Fully-Unbiased Model (Ürün Modu)

input\_text

Clear Submit

output

Flag

<https://huggingface.co/spaces/l2reg/Teknofest2023>

<https://github.com/L2-Regulasyon/Teknofest2023>

# DEMO

Spaces: l2reg/Teknofest2023

like 0

Running on T4

Logs

App

Files

Community

Settings

Yarışma Ekranı

Demo Ekranı

file

Drop File Here  
- or -  
Click to Upload

Clear

Submit

output

Flag

<https://huggingface.co/spaces/l2reg/Teknofest2023>

<https://github.com/L2-Regulasyon/Teknofest2023>

# ÇÖZÜMÜN KULLANILABİLİRLİĞİ

- Çözüm için eğitilmiş modeller doğrudan **HuggingFace** üzerinden indirilebilir.
- **GitHub** repository klonlanarak direk app.py üzerinden server ayağa kaldırılabilir.
- Modeller HuggingFace sınıfları kullanılarak oluşturulduğundan özel yazılmış başka bir inference koduna da kolay bir şekilde eklenebilir.
- Bütün model sınıfları kolaylıkla özelleştirilebilir ve başka bir verisetiyle yeniden eğitilebilir.
- Bütün rutinlerin çalışması için gereken ortam **Docker** formatında sunulmuştur.



# PROJENİN ÇIKTILARI

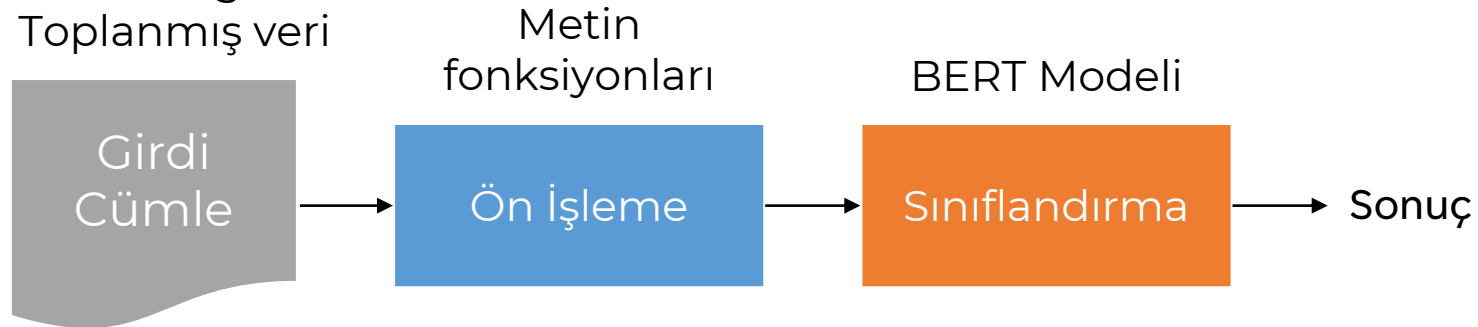
- Herkes tarafından kullanıma açık bir NLP sınıflandırma eğitim kiti oluşturulmuştur. Geleneksel ve transformer embedding yaklaşımlarını içermektedir.
- Yarışma hedefleri doğrultusunda eğitilen BERT modelleri herkese açık şekilde yayımlanmıştır.
- Modellerin eğitilmesi için gereken ön-işleme süreçleri kod ve döküman olarak paylaşılmıştır.
- Sürece ait her adımın yeniden nasıl çalıştırılabileceği raporlanmıştır.



# PROJENİN YOL HARİTASI

## MEVCUT YAPI

- Veritabanı
- Kullanıcı girdisi
- Toplanmış veri



# PROJENİN YOL HARİTASI

## POTANSİYEL SOSYAL ETKİ

Önerilen çözümün:

- İşitsel ve görsel medyada, eğitim çağındaki çocukların eğitim kaynağı olarak kullanabileceği materyallerin içerikleri metne dönüştürüldükten sonra ilgili etiketleri içerip içermediğini tespit etmede
- Müzik platformlarında, içeriklerin yaş sınırlaması tanımları muğlak, belirsiz olabilmektedir. Çocuklarının ilgili etiketleri barındıran müzikleri dinlemesini istemeyen ebeveynler için, çocuklarının dinleyebileceği türdeki müzikleri sınıflandırmada
- Devlet kurumlarında vatandaşlardan gelen istek/öneri/şikayetlerde; özel sektörde müşteri memnuniyeti artırmak için yapılan çalışmalarda metin dosyalarını sınıflandırma ve önceliklendirmede
- Çağrı merkezlerinde Speech2Text teknolojisi ile metne dökülen konuşmaları sınıflandırılarak operasyon verimliliğini artırmada

katkısı olabilir.



# TEŞEKKÜRLER