

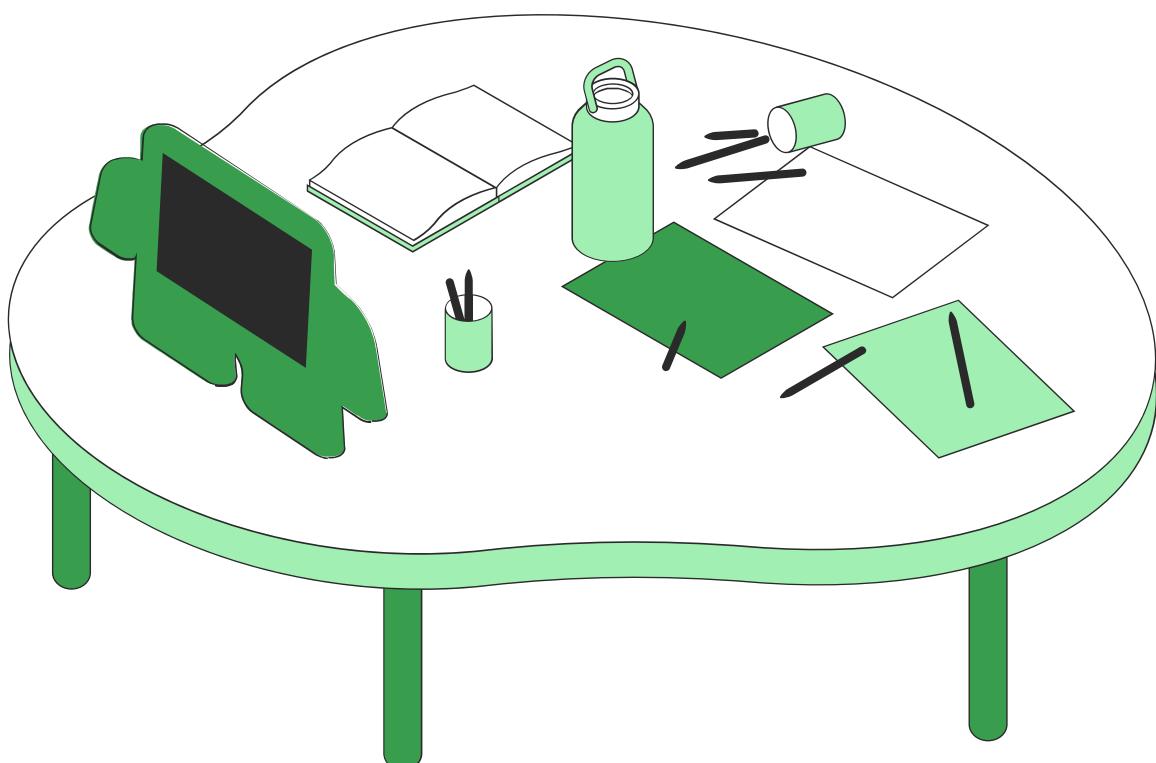
# Türkçe Nefret Söylemi Siniflandırması

HİVE HEROES

Berfin Duman  
Bike Sönmez



# iÇERİK



---

Projenin Motivasyonu

---

Projenin Amacı

---

Metodoloji

---

Araştırma- Uygulama- Geliştirme

---

Ürünleştime- İnovasyon

# Proje Motivasyonu

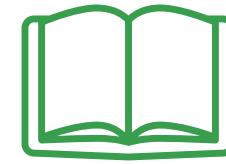


# Nefret söylemi nedir ?

Nefret söylemi veya nefretli konuşma, ırk , cinsiyet, yaş, ulus , din ya da cinsel yönelim gibi konulara dayanarak, belirli bir grup ya da kişiye yönelik nefret ifadesi içeren veya bu kişilere şiddet uygulanmasını teşvik eden aleni konuşmadır.



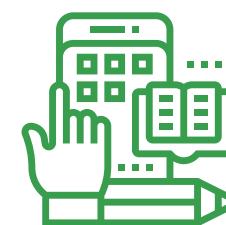
# Nefret Söylemini Önlemek ve Etkilerini Azaltmak İçin Ne Yapılabilir ?



Eğitim

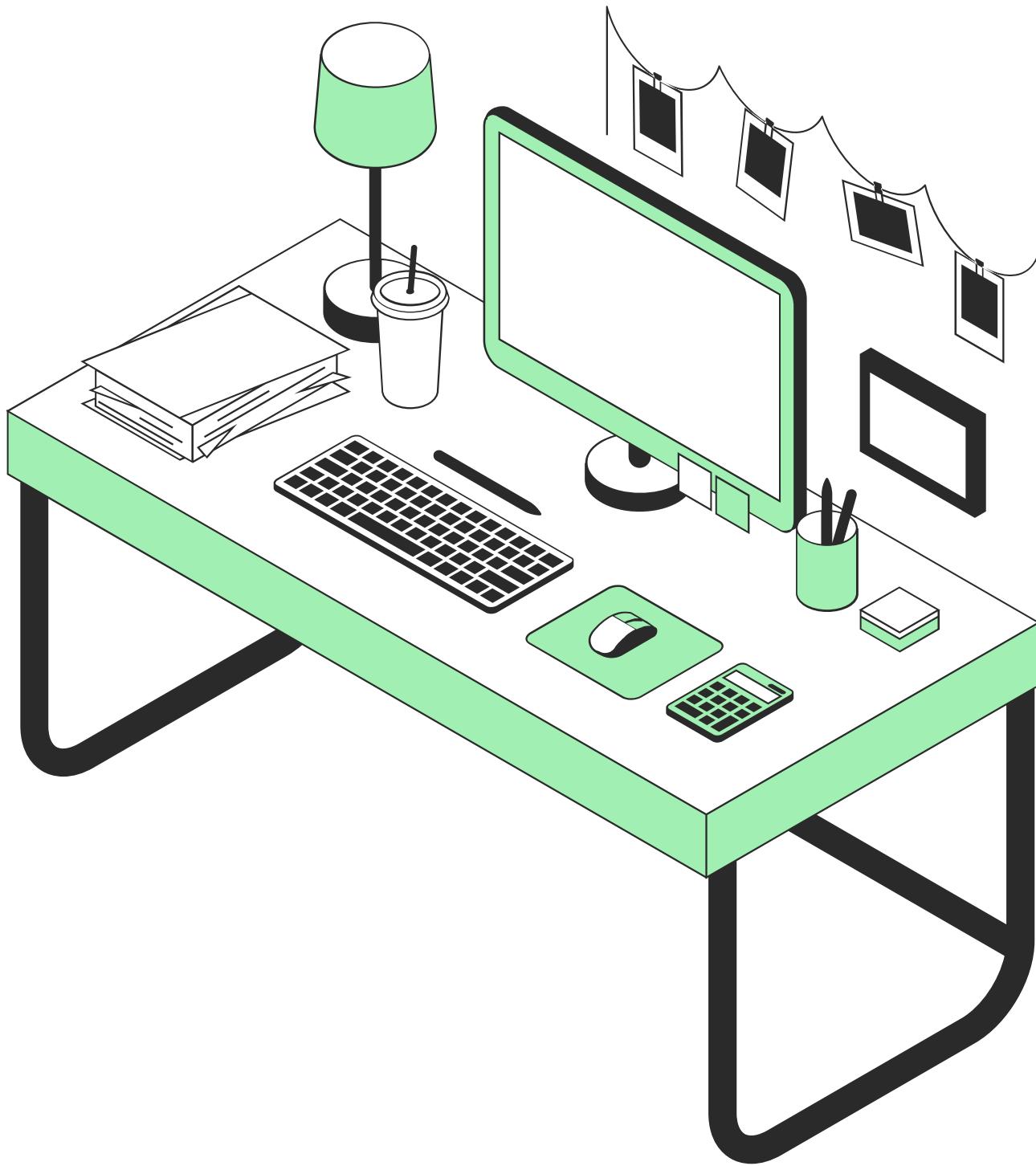


Kontrol Mekanizmaları



Teknolojik Uygulamaların  
Kullanımı

# Proje Amaci



# Metodoloji



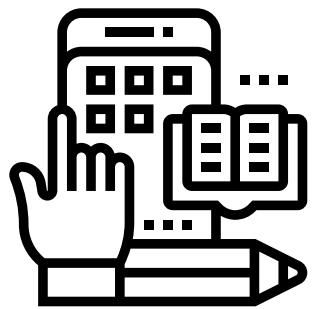
Araştırma



Analiz



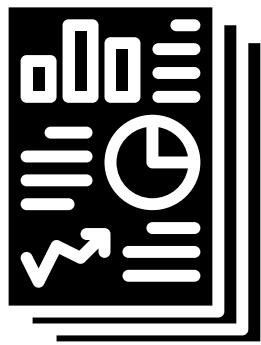
Uygulama



Geliştirme



Use case



# Araştırma



# Hangi Alanlarda Araştırma Yaptık ?

---

Nefret söylemi ve Etkileri

---

Geleneksel ve Gelişmiş NLP  
Yaklaşımları

---

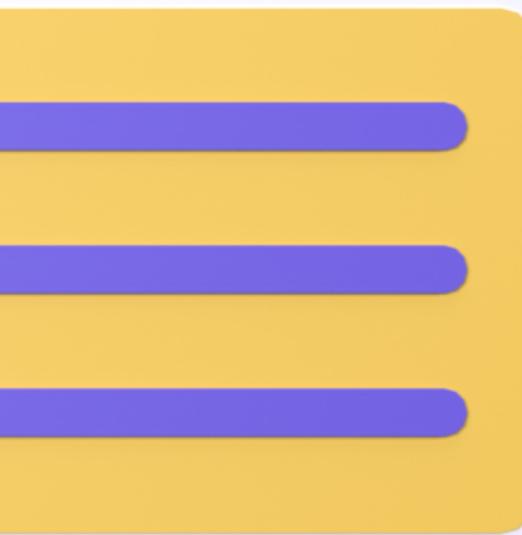
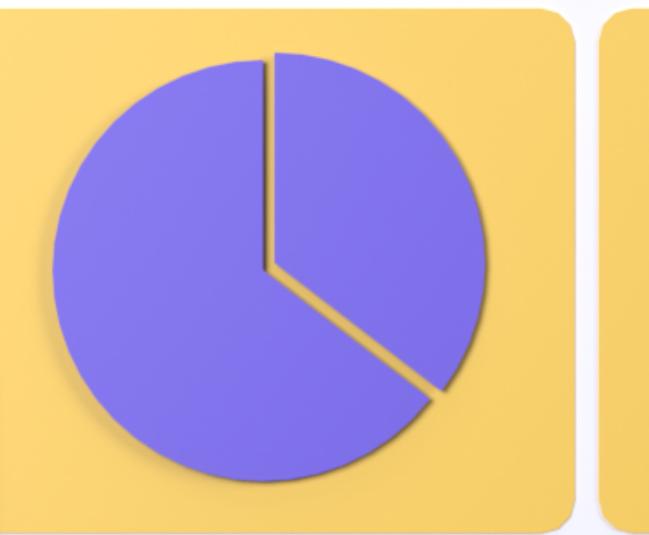
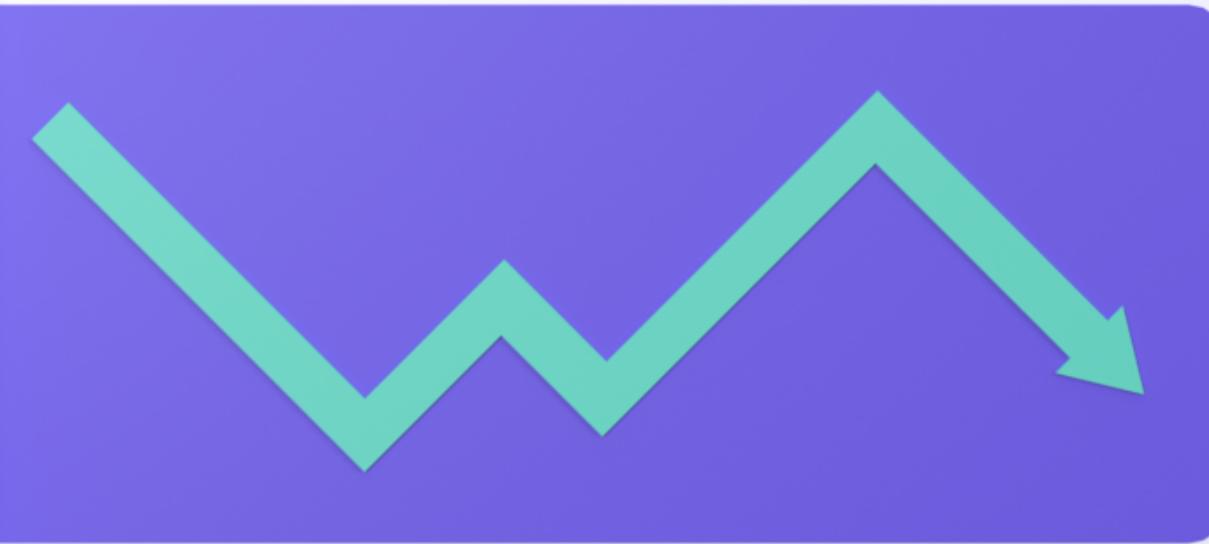
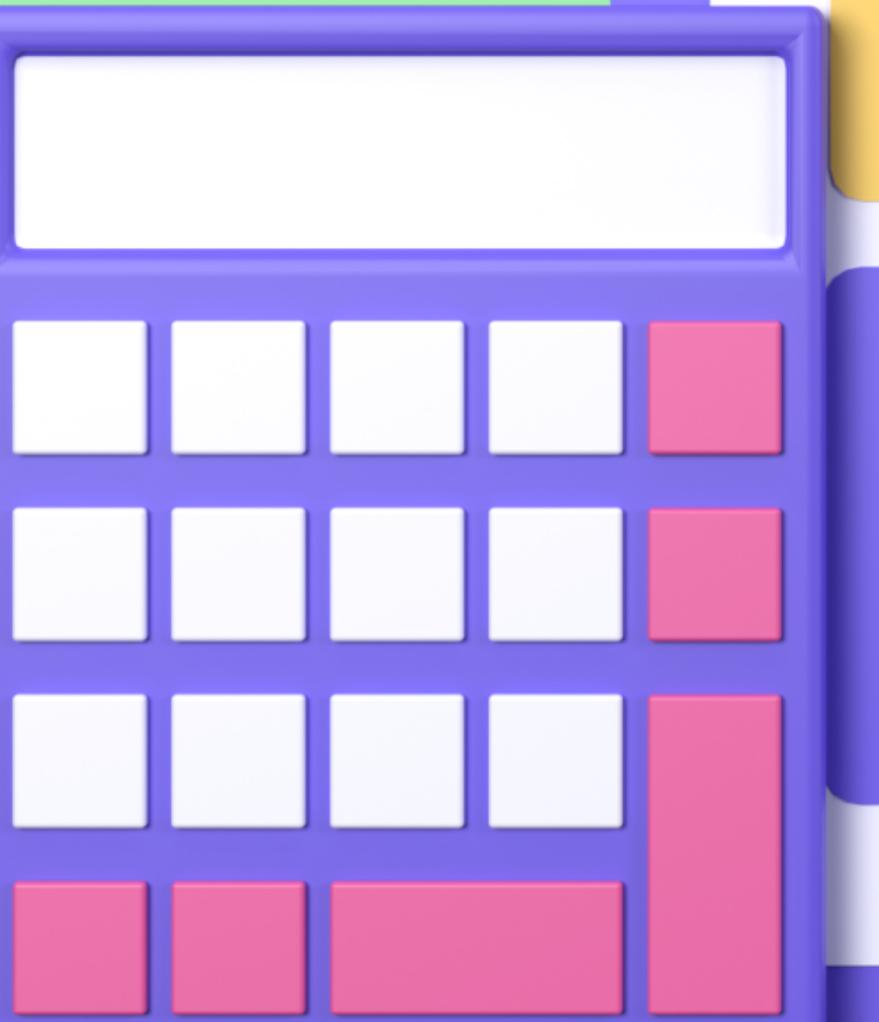
Türkçe NLP uygulamaları

---

Nefret Söylemi Tespitinin Sektörde  
Kullanılabileceği Alanlar

---

# Analiz



# Veri Analizi

Ne tür bir veri  
setine sahibiz



Veri setine  
hangi işlemleri  
uygulamalıyız



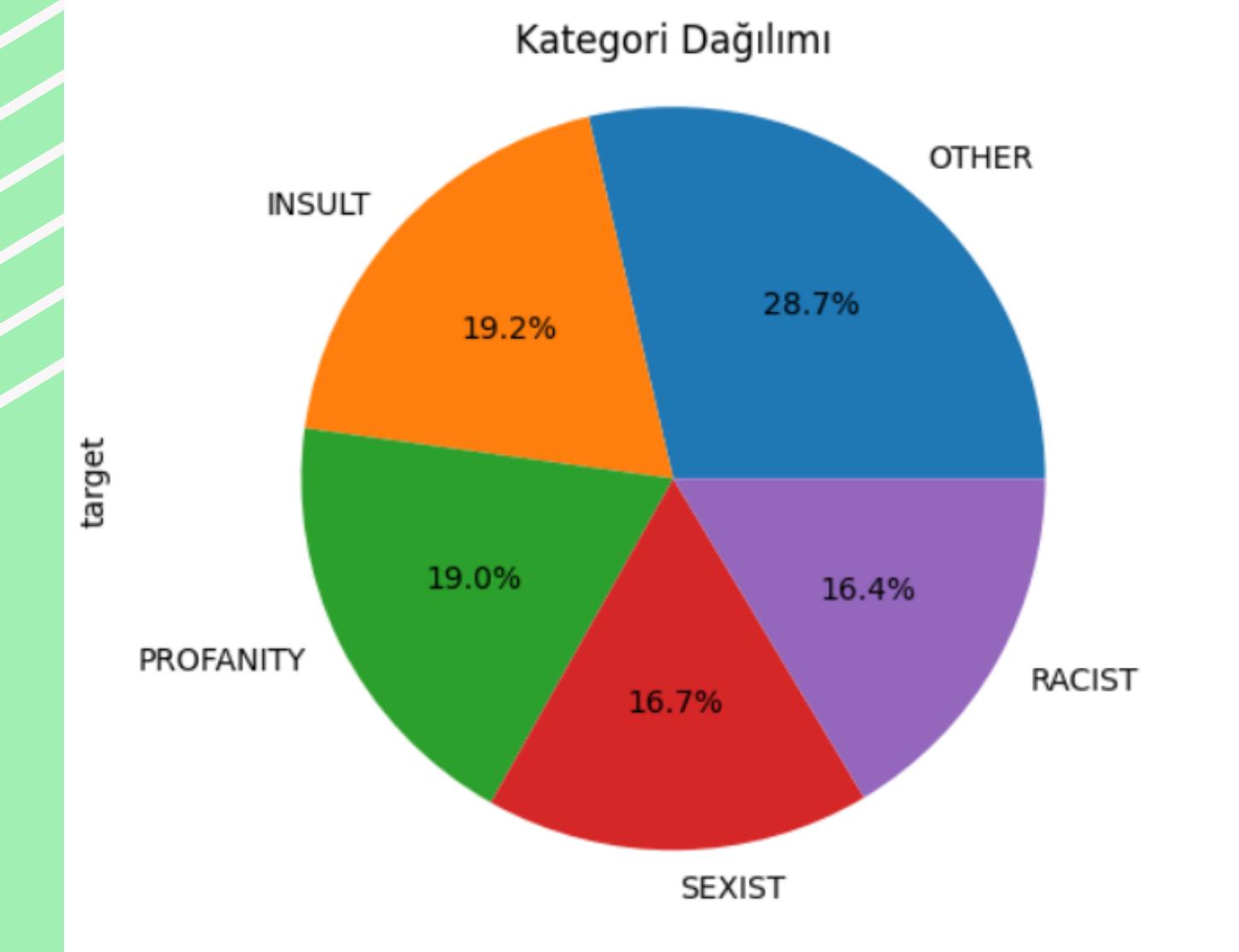
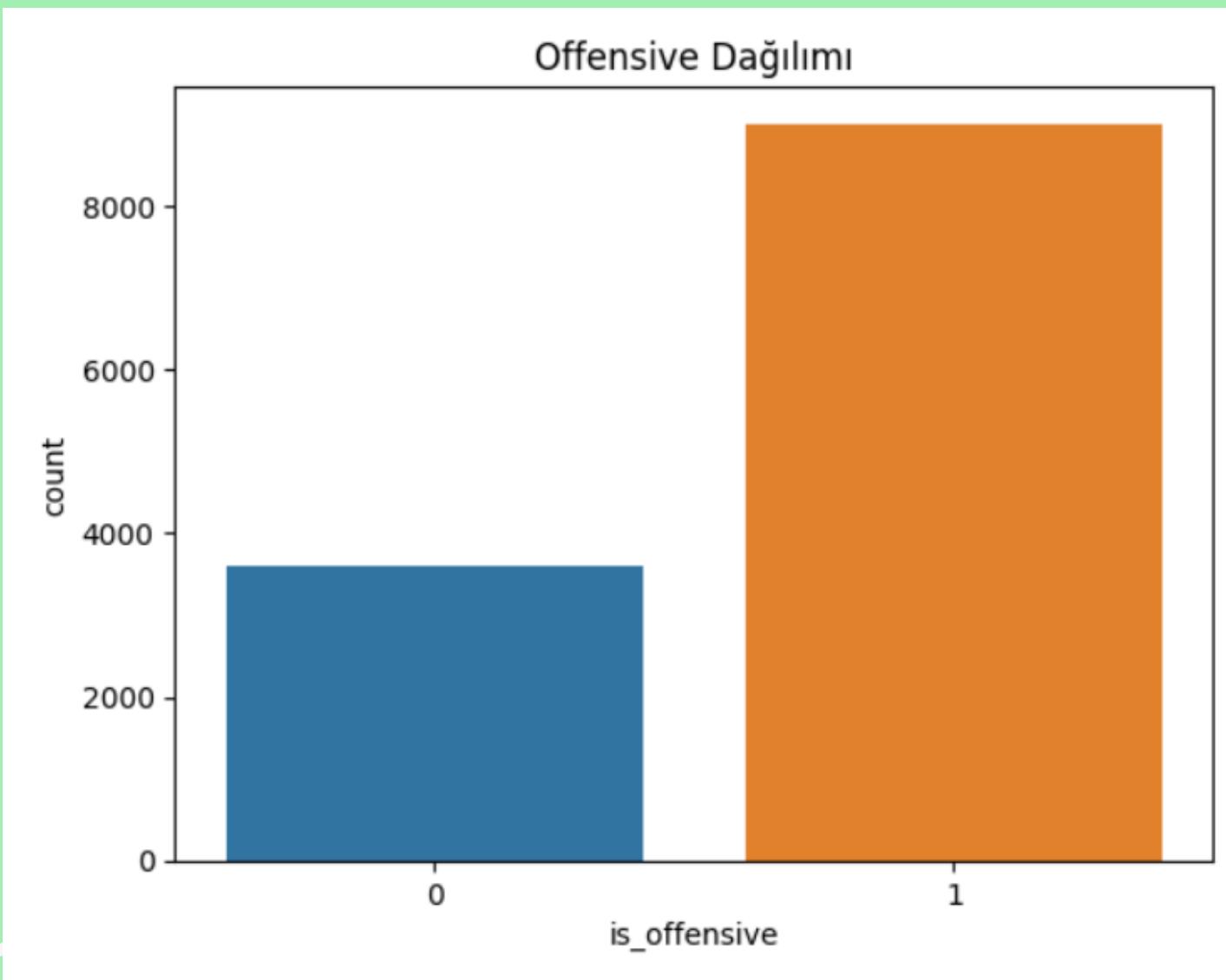
Veri setini  
modeller için  
nasıl hazır hale  
getirebiliriz



	<b><i>id</i></b>	<b><i>text</i></b>	<b><i>is_offensive</i></b>	<b><i>target</i></b>
<b>0</b>	81c11060-a240-4d54-841b-9e2916039e85	çürük dişli	1	INSULT
<b>1</b>	be80ebbf-b322-4c3b-afa1-94932ea80731	Bu adamın islama ve müslümanlara verdiği zarar...	1	RACIST
<b>2</b>	f99e2513-83ed-4076-ac72-b9e2cff3f049	erkekler zora gelmez	1	SEXIST
<b>3</b>	83ed2b2e-b815-4f36-9fc4-80a9050cf2d0	Utanmazın götüne kazık sokmuşlar bu tıkırtı ne...	1	PROFANITY
<b>4</b>	d93e05f7-bfdd-4cdb-99d8-3048761b30ff	otomasyon< sistemlerine= doğrudan bağlanabilir	0	OTHER
<b>5</b>	81eca520-15a0-4f5b-a1dd-5929bba34bdc	amını sikiğimin yarrağı	1	PROFANITY
<b>6</b>	a2497690-5201-4c28-bdb8-4206ea672765	ekipte bir cücemiz eksikti o da oldu	1	INSULT
<b>7</b>	b9122be9-5a9a-491e-8ecc-998ac05c6ef0	Kadın yönetici sayısı oldukça az olduğu için ş...	1	SEXIST
<b>8</b>	2dd1b0d1-45cd-4d4a-86c3-b290db5a0316	makara bunu duyunca avluya çıkışmış ve kızağa sa...	0	OTHER
<b>9</b>	d7781fd9-ee64-4b51-9abd-aa8eda638230	hay ebenin pembe patığine sıçayılm bok kafalı	1	PROFANITY

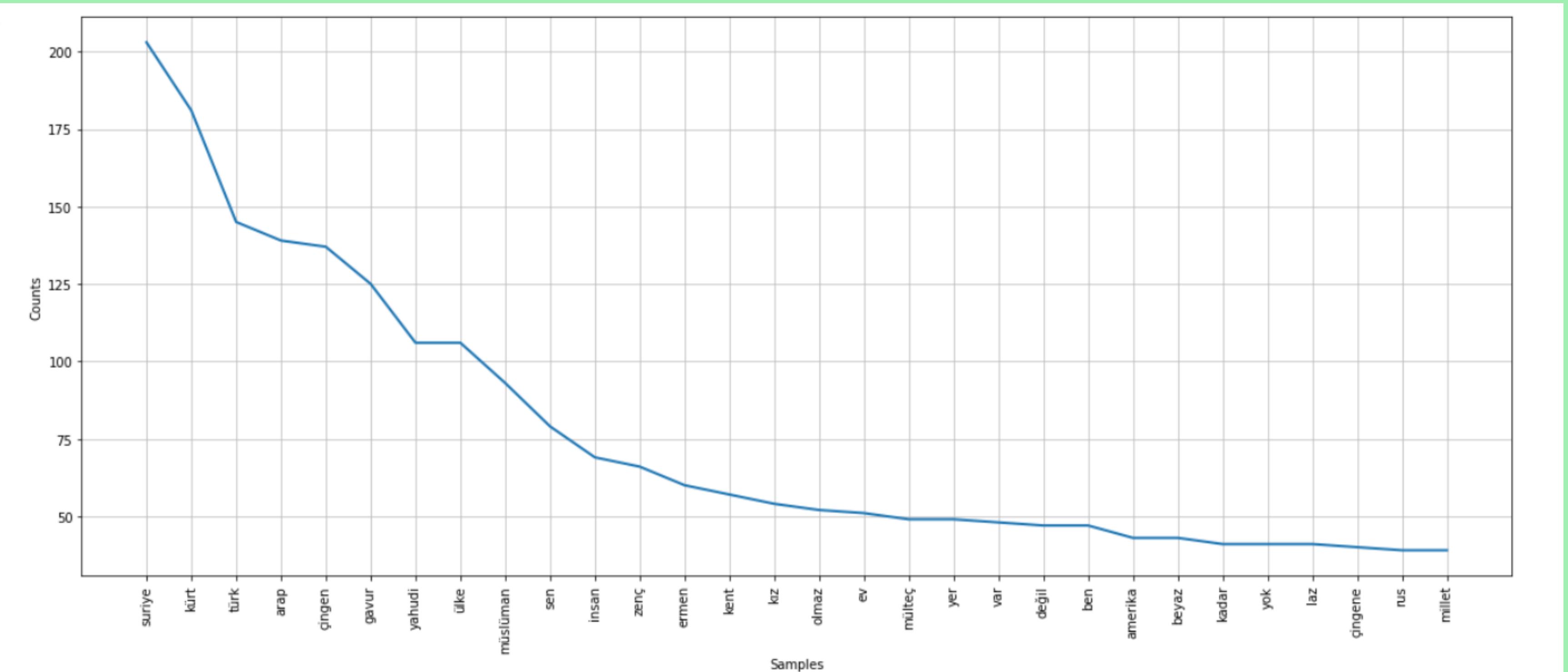
teknofest\_train\_final.csv

# Bazı Veri Dağılımları



# WordCloud ile Görselleştirme





Racist Kategorisine Ait Metinlerde En Çok Geçen Sözcükler

# Preprocessing

Özel karakterlerin silinmesi

---

Tokenizasyon

---

Durma Sözcüklerinin Silinmesi

---

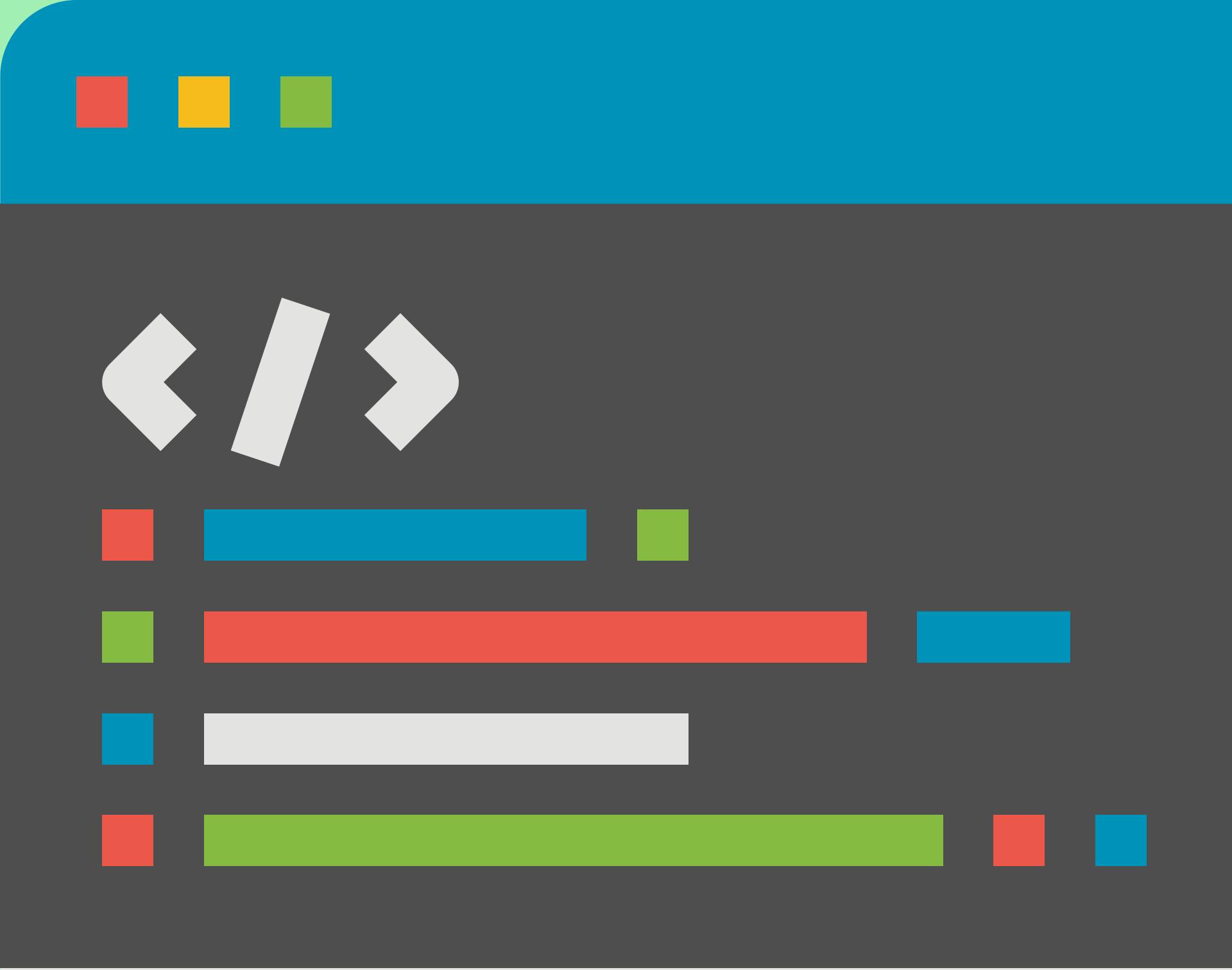
Lemmatizasyon

---

Vektörizasyon

---

# Uygulama



# Geleneksel ML Yaklaşımları



**Bernoulli Naive Bayes  
Classifier**



**Support Vector  
Machine (SVM)**



**Decision Tree  
Classifier**



**XGB Classifier**

# Dil Bazlı Gelişmiş NLP Yaklaşımları



**LSTM - BiLSTM**



**BERT**



**GPT**

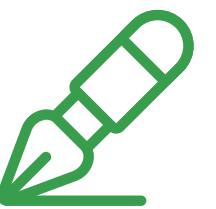


**Ensemble Modeller**

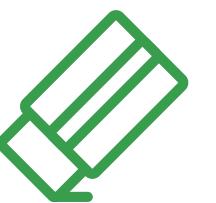
# Modellerde Karşılaştığımız Problemler Nelerdi ?



İnsult kategorisini ayırt  
etme problemi



Bazı modellerin Türkçe  
versiyonlarındaki hatalar

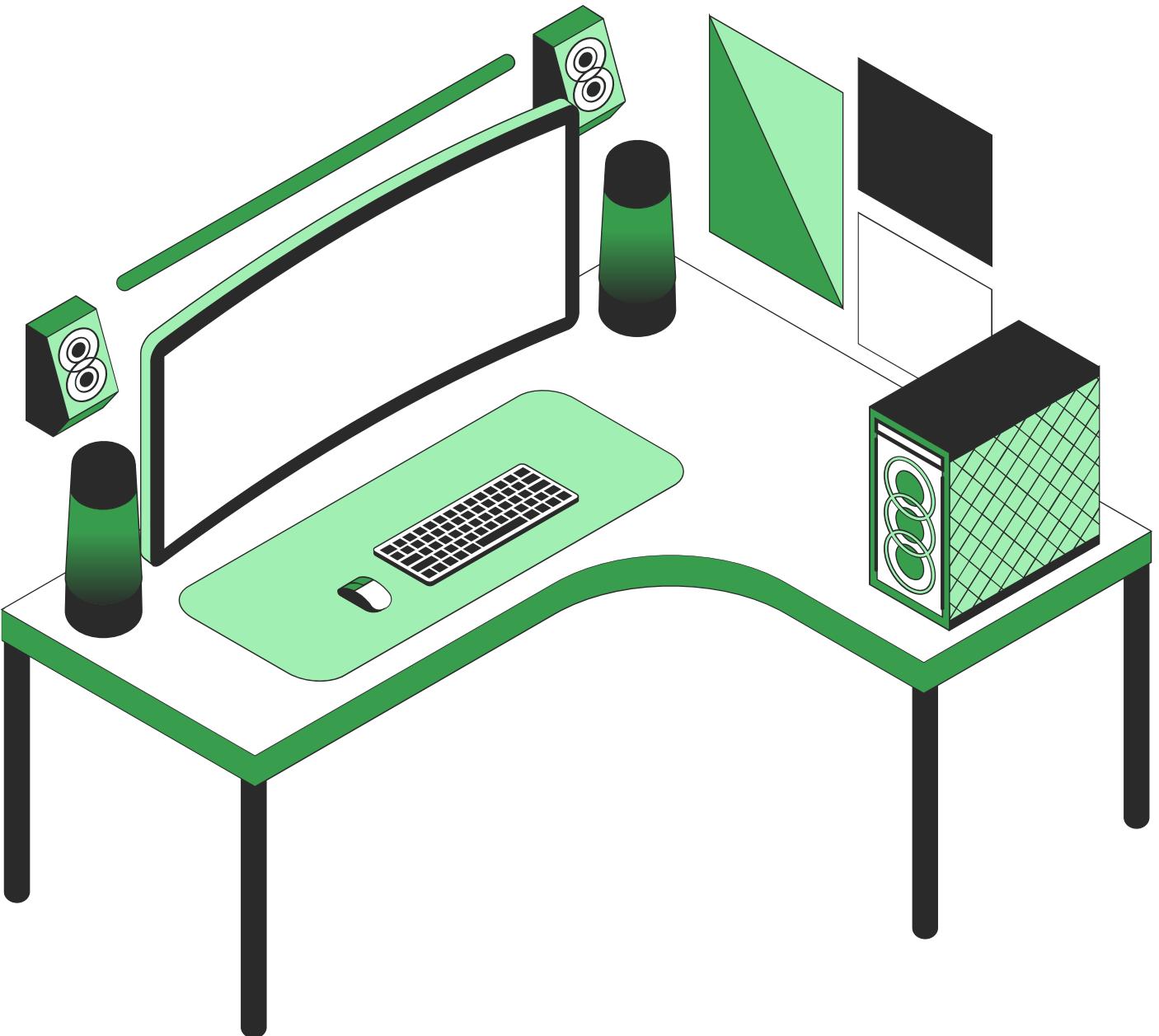


Model Hiperparametre  
Optimizasyonu

# Bizim Bu Problemlere Karşı Yaklaşımlarımız Nelerdi ?



- Insult kategorisindeki değerler için veri artırma
- Daha gelişmiş model kombinasyonları kullanma (Ensemble Modeller)
- Earlystopping ile en optimum epoch sayısını bulma (11)



**Neden  
Ensemble  
Model?**

# Ensemble Model Yaklaşımı



**BERT TÜRK**

- BERT- Multilingual
- BERT -Turkish
- Kelime dizileri üzerinde çalışır



**BİLSTM**

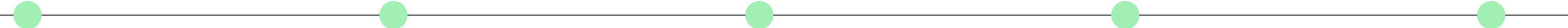
- Sıralı metinleri sınıflama



**Convolutional neural network(CNN)**

- Metindeki desenleri tanımak için
- Karakter bazlı embedding

# Neler Yaptık



**Geleneksel  
Sınıflandırma  
Modelleri**

**Transformer  
Modeller**

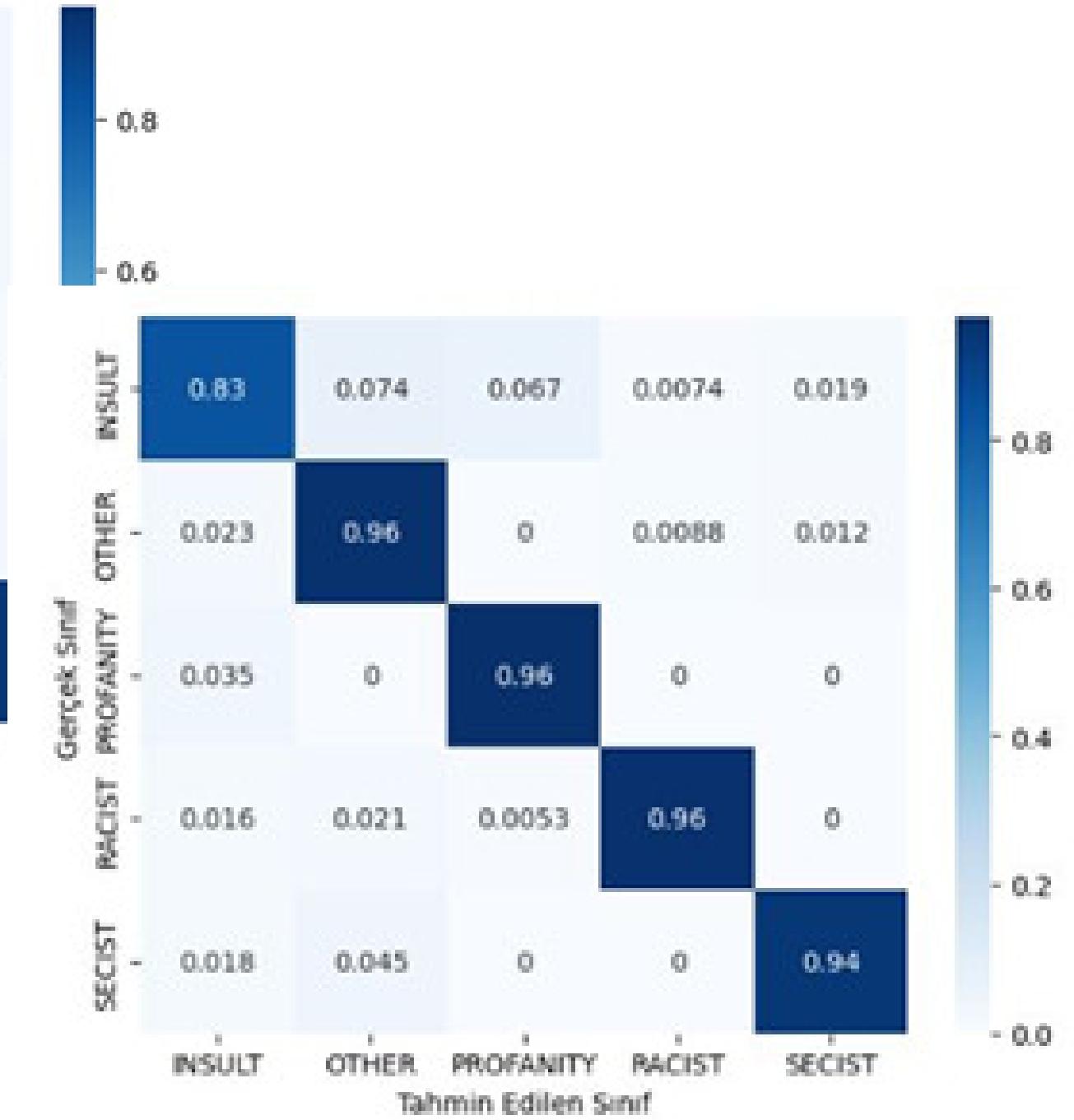
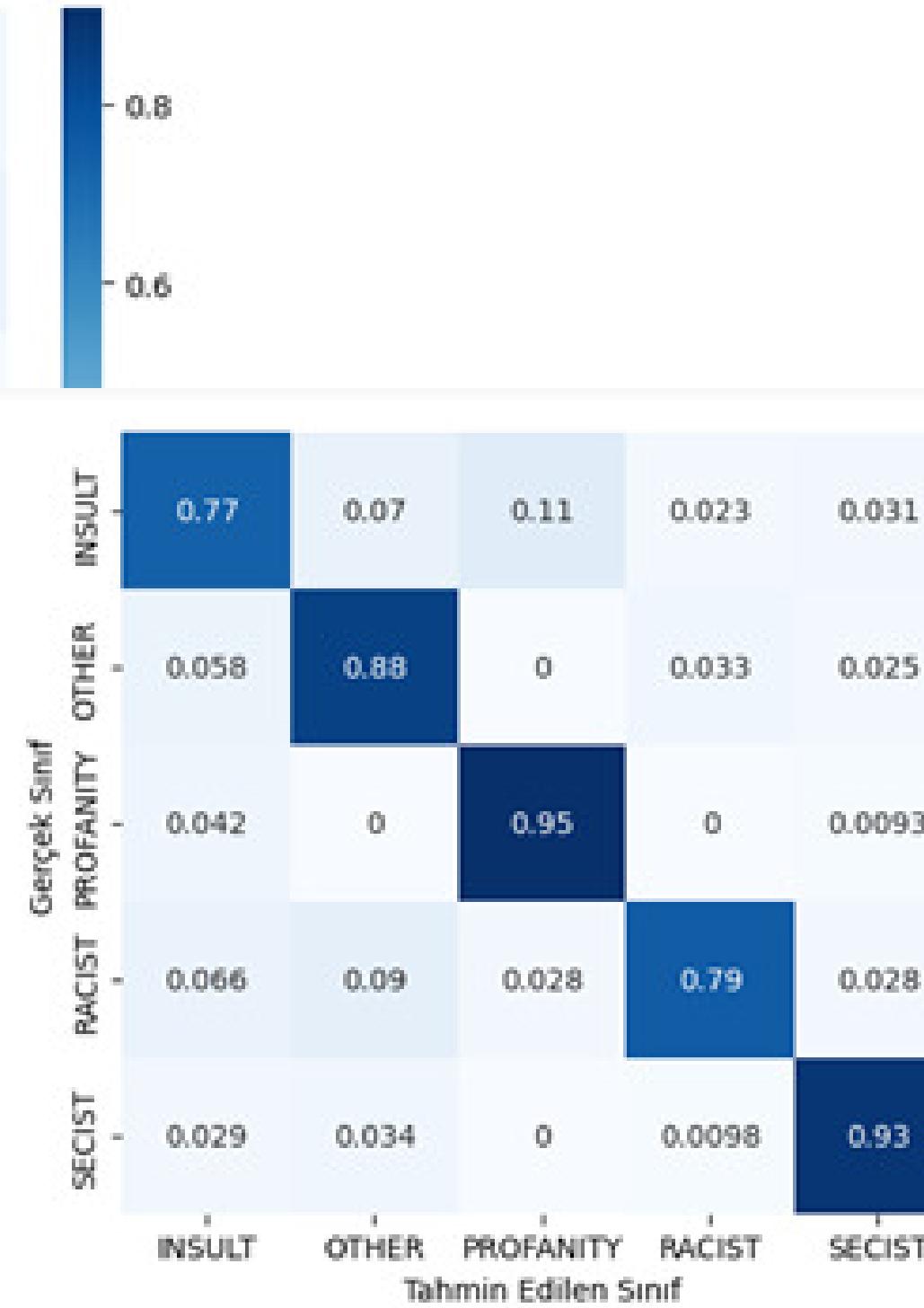
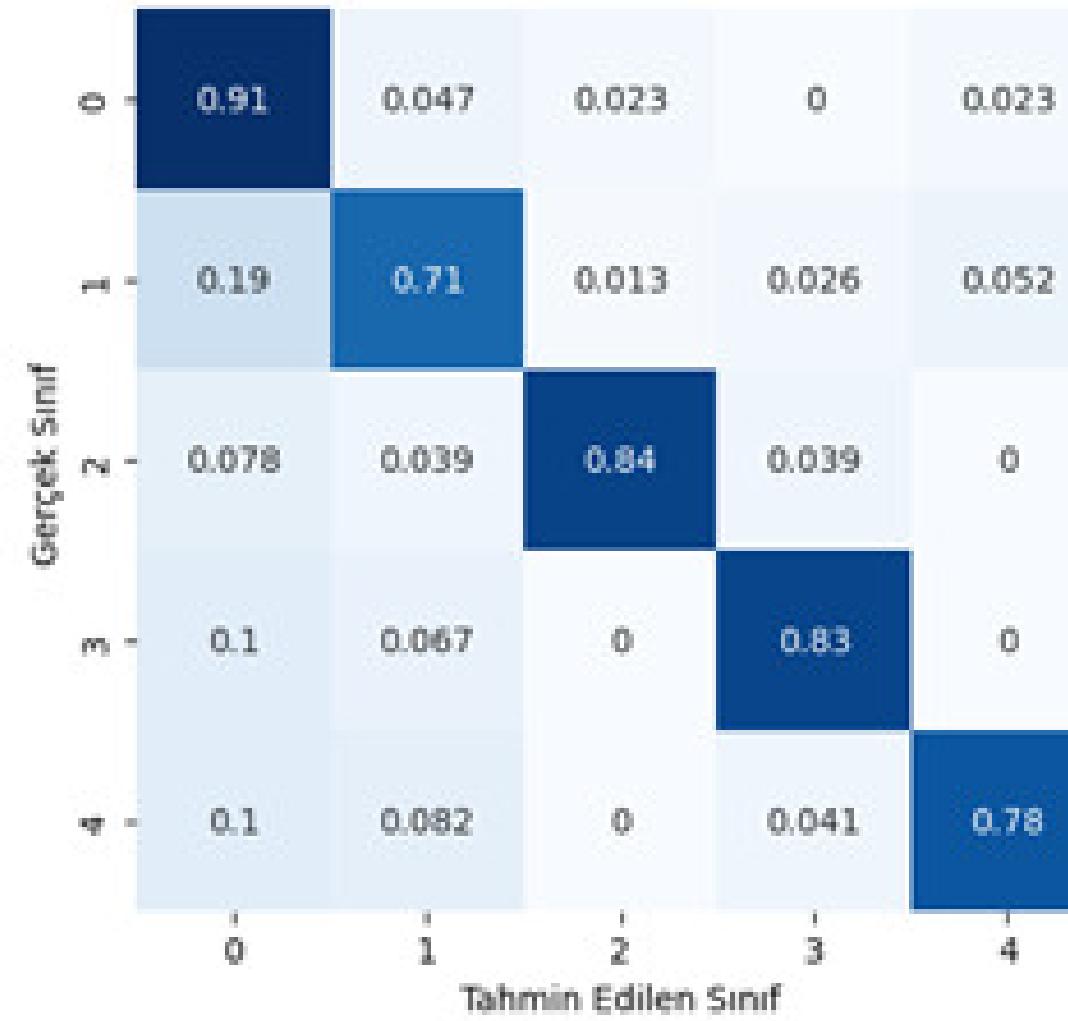
**BERT-  
Multilingual**

**BERT-  
Türkçe**

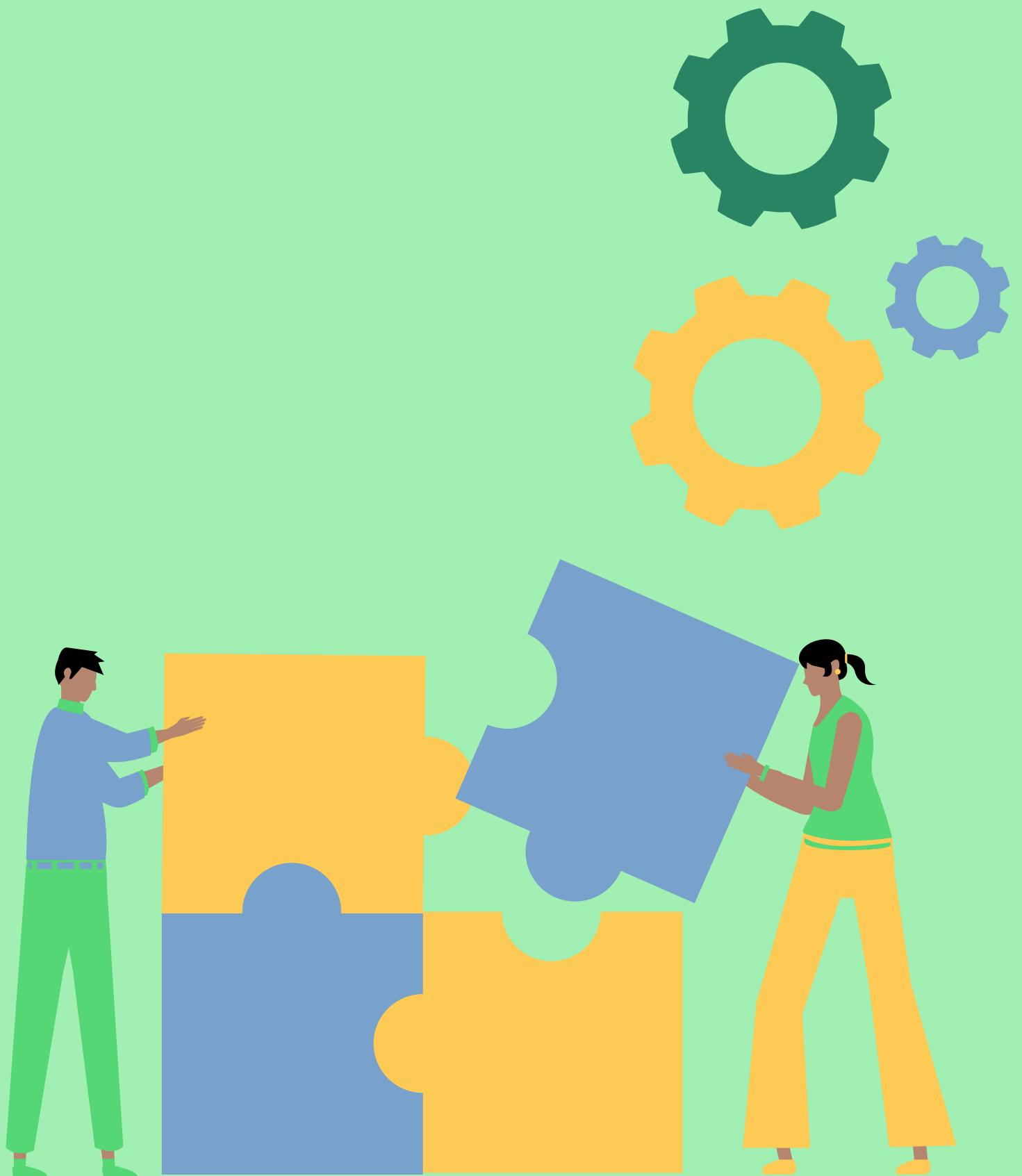
**Ensemble  
yaklaşım**



# Skorlar



# Elde Ettiğimiz Çıktıyı Nasıl Bir Ürüne Dönüştürebiliriz ?



Telco ve OTT  
Sektörlerinde Nasıl  
Kullanabilirimiz ?



# Use Case Modeli





Çocuklara Özel Dijital  
Ortam

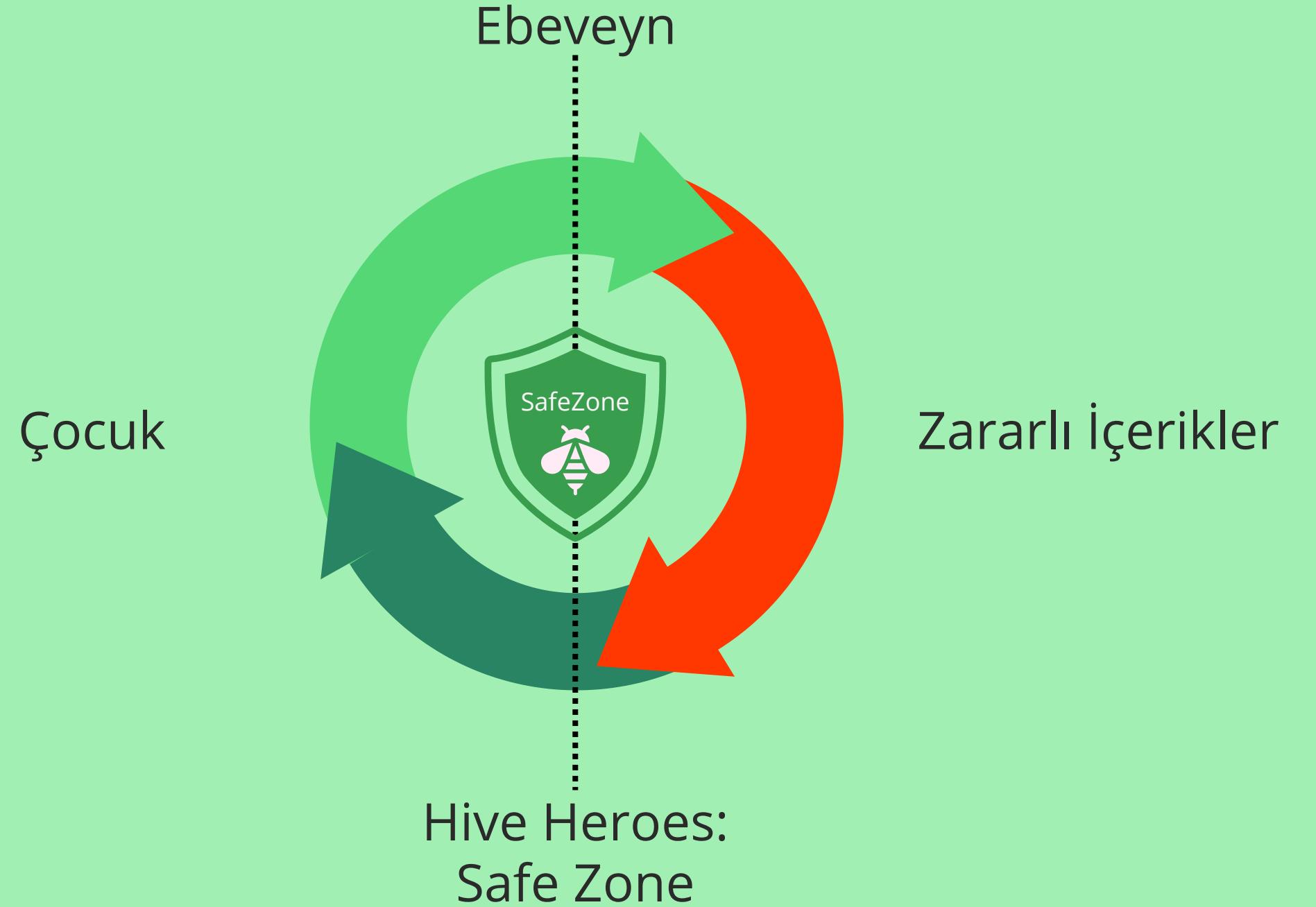


Canlı Yayınlar İçin  
Eklentiler



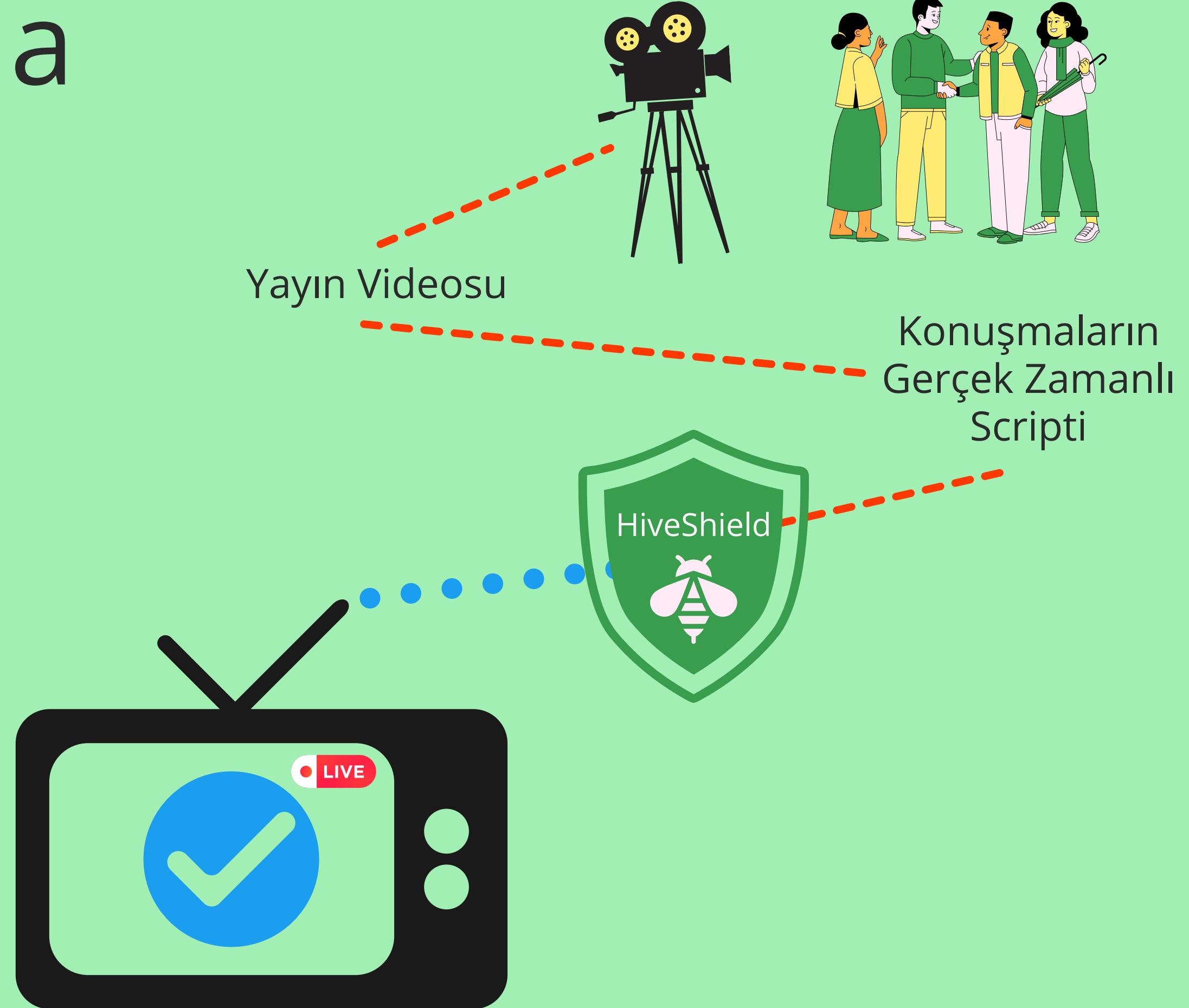
Sosyal Medya  
Filtrelemesi

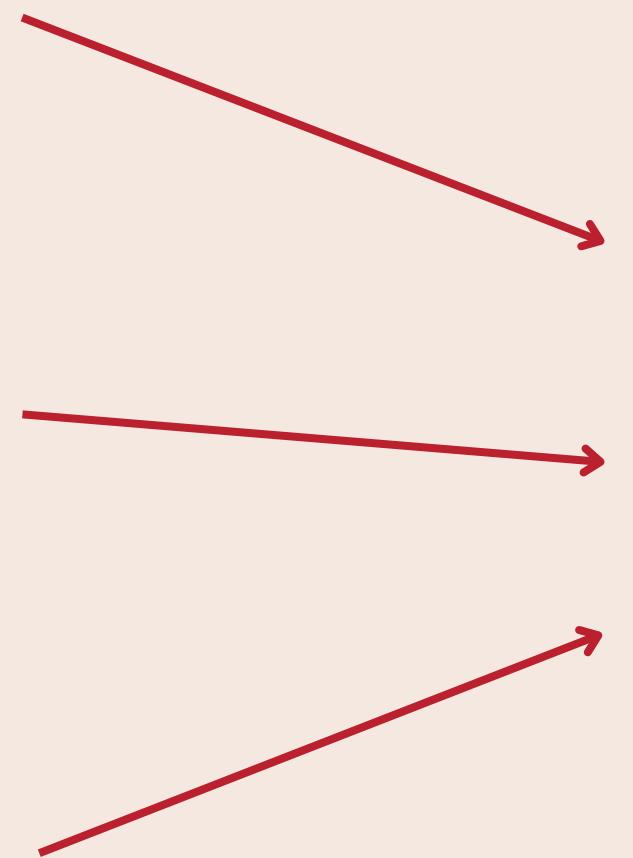
# Çocukları Korumak İçin



Çocuklara Özel Dijital  
Ortam Sağlayan Program

# Canlı yayınılarda insanların ne konuşacağını bilemeyiz...





# Bizi Dinlediğiniz İçin Teşekkürler

HiveHeroes

Berfin Duman  
berfiinduman@gmail.com

Bike Sönmez  
bikesnmz@outlook.com

