

Ödev No:2

Ödev Konusu: Sorudan cevaba, Hangisi iyi

https://docs.google.com/spreadsheets/d/1Woh-A5oTJ715ivgIsu6NCkdav_k46igr/edit?usp=drive_link&ouid=103534328876075299488&rtopof=true&sd=true adresinde bulunan dosyadaki soru, gpt4o cevabı, deepseek cevabı ve hangisi iyi (1,2,3,4) bilgilerini kullanarak

A) Sorudan cevaba (5 puan)

Rasgele 1000 soru/satır seçip bu sorular üzerinde:

Bir soruya en benzeyen 5 gpt4o cevabını bulup gerçek cevaba göre topl ve topl başarılarını hesaplamak.

Bir soruya en benzeyen 5 deepseek cevabını bulup gerçek cevaba göre topl ve topl başarılarını hesaplamak.

topl ve topl başarıları ile 1 2 3 4 sınıflarının korelasyonu üzerinde yorum yapmak, örnek bazında ve toplu olarak **(1 puan)**

B) Hangisi iyi (5 puan)

s: soru, g: gpt4o, d: deepseek vektörlerini göstermek üzere

s, g, d, s-g, s-d, g-d, |s-g|, |s-g|-|s-d| vb. vektörleri giriş olarak alıp 1 2 3 4 sınıflarını tahmin eden bir model eğitmek

veri kümesini tamamının (yaklaşık 13 bin) %80 eğitim %20 test olarak kullanınız.

Farklı vektör kombinasyonlarının sınıflandırma başarılarını karşılaştırmak

Metinleri vektörlere dönüştürmek için aşağıdaki 3 modeli kullanınız:

e5: <https://huggingface.co/intfloat/multilingual-e5-large-instruct>

cosmosE5: <https://huggingface.co/ytu-ce-cosmos/turkish-e5-large>

jina: <https://huggingface.co/jinaai/jina-embeddings-v3>

Bu modellerin A'da topl ve topl başarıları, B'de sınıflandırma performansları üzerindeki etkilerini ölçünüz.

Bulgularınızı tablolar ve grafiklerle sunun, yorumlayın.

Ödevin Son Teslim Tarihi: 29 Nisan 2025 saat 09:30

Ödevin Teslim Şekli: online.yildiz.edu.tr

Bu ödevde en çok 2 kişilik gruplar halinde çalışacaktır.

Ödevde Teslim Edilecekler:

1-Ödev Raporu (Bulgularınızı açıklayan tablo ve grafikler, Bulgularınıza dair yorumlarınız) (%40)

2-Programın çalıştırılabilir hali ve açıklama içeren kodları (değişkenlerin ne için kullanıldıkları, algoritmanın adımları) (%30)

3-Raporunuzu anlattığınız (5-8 dakika uzunluğunda) ve youtube'a yüklenmiş videonun adresi (%30)