



BLM3051 - Yapay Zeka

Ödev No:2

Öğrenci Adı: İclal ERTÜRK

Öğrenci Numarası: 21011037

Dersin Eğitmeni: M. Fatih Amasyalı

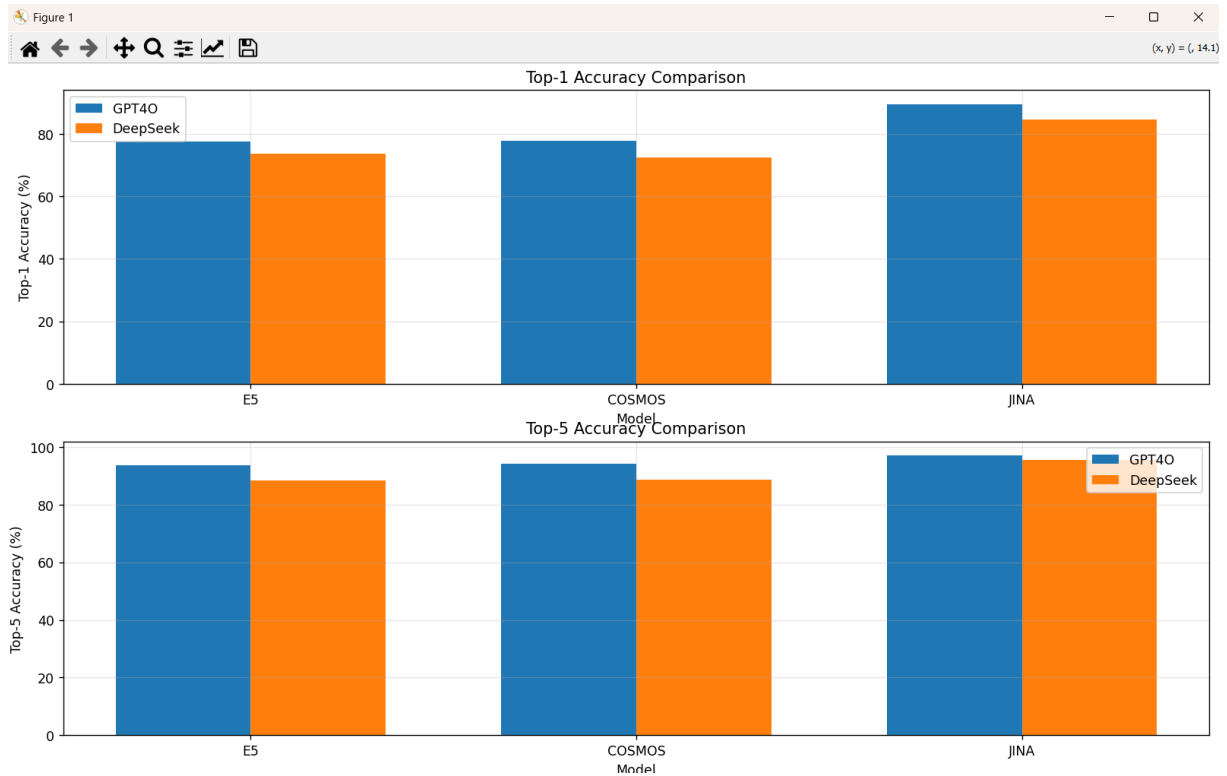
Video Linki: <https://youtu.be/M2OuYWY4Jxk>

İçindekiler

| | |
|--|----|
| A-Sorudan Cevaba | 3 |
| A-Sorudan Cevaba: top1 ve top5 başarıları ile 1 2 3 4 sınıflarının korelasyonu üzerinde yorum yapmak, örnek bazında ve toplu olarak..... | 4 |
| E5 -Gpt4o..... | 5 |
| E5-Deepseek..... | 6 |
| Cosmos-Gpt4o | 7 |
| Cosmos-Deepseek | 8 |
| Jina-Gpt4o | 9 |
| Jina- Deepseek..... | 10 |
| Elde Edilen Sonuçların ve Grafiklerin Yorumlanması | 11 |
| B-Hangisi İyi | 12 |
| Sonuçların Yorumlanması | 13 |

A-Sorudan Cevaba

Ödevin A şıkkı için 'binOrnekSecme.py' dosyasının içinde veri setinden rastgele 1000 soru seçilmiş ve 'bin_rastgele_soru.xlsx' dosyası içerisine cevapları ile beraber kaydedilmiştir. Seçilen soruların her model için embedding çıkarımı 'embedding1000icin.py' dosyası içerisinde gerçekleştirilmiş ve çıkarılan embeddingler 'embedding1000' klasoru içerisine modelAdı_bin_rastgele_soru_s formatında kaydedilmiştir (soru için s, deepseek için d, gpt için g). Daha sonra 'a_sorudan_cevaba.py' dosyası içerisinde top1 ve top5 başarılarını hesaplamak için 'calculate_topk' fonksiyonu yazılmış ve bu fonksiyon içerisinde her soru için o an incelenen yapay zeka türünün tüm cevapları ile benzerliklerine cosinüs benzerliği kullanılarak bakılmıştır. Soruların doğru cevabı ise o soru ile aynı indise sahip olan cevaptır. Bulunan cosinüs benzerlikleri sıralanmış ve doğru cevap ilk 5 içerisinde mi diye bakılmıştır eğer ilk 5 içerisindeyse top5 değişkeninin değeri artırılmış ve ilk 1 de mi diye bakılmıştır. Eğer ilk 1 de ise top1 değeri artırılmıştır. Bu işlem tüm sorular için tekrarlanmıştır. Fonksiyonun sonunda ise top5/n ve top1/n değerleri başarı oranları olarak döndürülmüştür. evulate_model fonksiyonu içerisinde tüm modeller ve yapay zeka türleri için bu fonksiyon çağırılmış ve sonuçlar görselleştirilmiştir.



Bu sonuçlar incelendiğinde modellerin top5 başarısının top1 başarısından daha yüksek olduğu, 'Jina' modelinin genel olarak daha iyi sonuçlar verdiği ve modellerin 'GPT4O' için daha iyi sonuçlar verdiği gözlemlenmiştir. Bu da GPT4O'nun soruları ve cevapları arasında daha güçlü bir semantik ilişki olduğunu ve Jina'nın embedding'lerinin soru-cevap ilişkisini daha iyi yakaladığını göstermektedir.

```
--- E5 modeli için değerlendirme başlıyor ---
Top-K hesaplanıyor: 100%|████████████████████| 1000/1000 [00:03<00:00, 277.76it/s]
E5 - GPT40 - Top1: 77.60%, Top5: 93.60%
Top-K hesaplanıyor: 100%|████████████████████| 1000/1000 [00:03<00:00, 259.12it/s]
E5 - DeepSeek - Top1: 73.60%, Top5: 88.30%

--- COSMOS modeli için değerlendirme başlıyor ---
Top-K hesaplanıyor: 100%|████████████████████| 1000/1000 [00:03<00:00, 276.86it/s]
COSMOS - GPT40 - Top1: 77.70%, Top5: 94.10%
Top-K hesaplanıyor: 100%|████████████████████| 1000/1000 [00:03<00:00, 286.63it/s]
COSMOS - DeepSeek - Top1: 72.40%, Top5: 88.70%

--- JINA modeli için değerlendirme başlıyor ---
Top-K hesaplanıyor: 100%|████████████████████| 1000/1000 [00:03<00:00, 306.14it/s]
JINA - GPT40 - Top1: 89.50%, Top5: 97.10%
Top-K hesaplanıyor: 100%|████████████████████| 1000/1000 [00:03<00:00, 308.91it/s]
JINA - DeepSeek - Top1: 84.50%, Top5: 95.50%

Results Table:
Model  GPT40 Top-1  GPT40 Top-5  DeepSeek Top-1  DeepSeek Top-5
E5      77.6    93.6      73.6    88.3
COSMOS  77.7    94.1      72.4    88.7
JINA    89.5    97.1      84.5    95.5
```

A-Sorudan Cevaba: top1 ve top5 başarıları ile 1 2 3 4 sınıflarının korelasyonu üzerinde yorum yapmak, örnek bazında ve toplu olarak

Bu istenenler için 'aKorelasyon.py' dosyası içerisinde anayze_corrolation fonksiyonu ile örnek bazlı korelasyon, sınıf bazında Top-1 ve Top-5 doğruluk oranları, ve sınıf bazlı korelasyon değerleri analiz edilmiştir. Analizler için koddaki cosine benzerliği, Pearson ve Spearman korelasyonları kullanılmıştır.

Aşağıda Elde edilen sonuçlar ve grafiklere yer verilmiştir. Akabinde ise bu sonuçlar yorumlanmıştır.

E5-Gpt4o

e5 - GPT4O Korelasyon Sonuçları:

Top-1 Başarı Oranı: 77.60%

Top-5 Başarı Oranı: 93.60%

Top-1 Pearson Korelasyonu: 0.529

Top-5 Pearson Korelasyonu: 0.512

Top-1 Spearman Korelasyonu: 0.510

Top-5 Spearman Korelasyonu: 0.404

Örnek Bazlı Sonuçlar (İlk 10 örnek):

Örnek | Sınıf | Top-1 | Top-5 | Benzerlik Skoru

| | | | | |
|---|---|---|---|--------|
| 0 | 2 | 1 | 1 | 0.8808 |
| 1 | 1 | 0 | 0 | 0.8354 |
| 2 | 3 | 1 | 1 | 0.8958 |
| 3 | 3 | 0 | 1 | 0.8581 |
| 4 | 2 | 1 | 1 | 0.8976 |
| 5 | 3 | 1 | 1 | 0.9525 |
| 6 | 4 | 1 | 1 | 0.9284 |
| 7 | 1 | 1 | 1 | 0.9174 |
| 8 | 1 | 1 | 1 | 0.8796 |
| 9 | 3 | 1 | 1 | 0.8803 |

Sınıf Bazlı Başarı Oranları:

Sınıf 1:

Top-1 Başarı Oranı: 79.46%

Top-5 Başarı Oranı: 94.64%

Sınıf 2:

Top-1 Başarı Oranı: 77.02%

Top-5 Başarı Oranı: 95.32%

Sınıf 3:

Top-1 Başarı Oranı: 77.44%

Top-5 Başarı Oranı: 92.68%

Sınıf 4:

Top-1 Başarı Oranı: 73.47%

Top-5 Başarı Oranı: 89.80%

Sınıf Bazlı Korelasyonlar:

Sınıf 1:

Top-1 Pearson Korelasyonu: 0.515

Top-5 Pearson Korelasyonu: 0.471

Top-1 Spearman Korelasyonu: 0.479

Top-5 Spearman Korelasyonu: 0.367

Sınıf 2:

Top-1 Pearson Korelasyonu: 0.483

Top-5 Pearson Korelasyonu: 0.430

Top-1 Spearman Korelasyonu: 0.476

Top-5 Spearman Korelasyonu: 0.339

Sınıf 3:

Top-1 Pearson Korelasyonu: 0.541

Top-5 Pearson Korelasyonu: 0.542

Top-1 Spearman Korelasyonu: 0.519

Top-5 Spearman Korelasyonu: 0.434

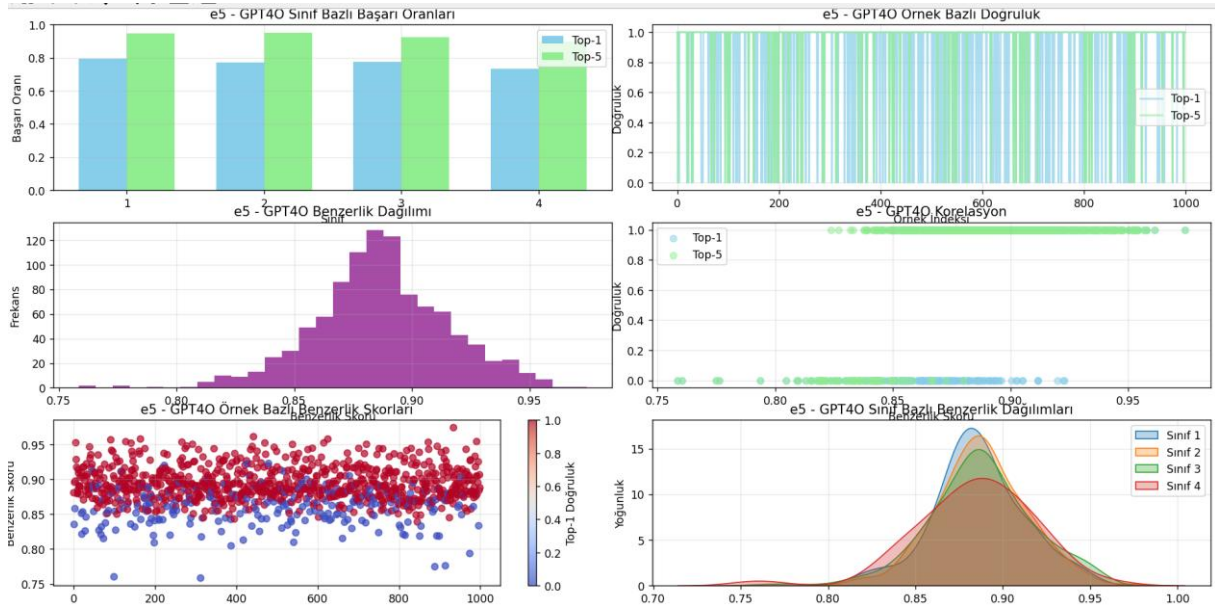
Sınıf 4:

Top-1 Pearson Korelasyonu: 0.681

Top-5 Pearson Korelasyonu: 0.606

Top-1 Spearman Korelasyonu: 0.680

Top-5 Spearman Korelasyonu: 0.496



E5-Deepseek

e5 - DeepSeek Korelasyon Sonuçları:

Top-1 Başarı Oranı: 73.60%

Top-5 Başarı Oranı: 88.30%

Top-1 Pearson Korelasyonu: 0.622

Top-5 Pearson Korelasyonu: 0.591

Top-1 Spearman Korelasyonu: 0.604

Top-5 Spearman Korelasyonu: 0.514

Örnek Bazlı Sonuçlar (İlk 10 örnek):

Örnek | Sınıf | Top-1 | Top-5 | Benzerlik Skoru

| | | | | |
|---|---|---|---|--------|
| 0 | 2 | 1 | 1 | 0.8711 |
| 1 | 1 | 0 | 0 | 0.8280 |
| 2 | 3 | 1 | 1 | 0.8911 |
| 3 | 3 | 0 | 1 | 0.8712 |
| 4 | 2 | 1 | 1 | 0.9100 |
| 5 | 3 | 1 | 1 | 0.9100 |
| 6 | 4 | 1 | 1 | 0.9322 |
| 7 | 1 | 1 | 1 | 0.8521 |
| 8 | 1 | 1 | 1 | 0.8654 |
| 9 | 3 | 0 | 0 | 0.7959 |

Sınıf Bazlı Başarı Oranları:

Sınıf 1:

Top-1 Başarı Oranı: 71.43%

Top-5 Başarı Oranı: 87.05%

Sınıf 2:

Top-1 Başarı Oranı: 75.32%

Top-5 Başarı Oranı: 88.51%

Sınıf 3:

Top-1 Başarı Oranı: 72.97%

Top-5 Başarı Oranı: 88.41%

Sınıf 4:

Top-1 Başarı Oranı: 81.63%

Top-5 Başarı Oranı: 91.84%

Sınıf Bazlı Korelasyonlar:

Sınıf 1:

Top-1 Pearson Korelasyonu: 0.612

Top-5 Pearson Korelasyonu: 0.631

Top-1 Spearman Korelasyonu: 0.588

Top-5 Spearman Korelasyonu: 0.541

Sınıf 2:

Top-1 Pearson Korelasyonu: 0.649

Top-5 Pearson Korelasyonu: 0.588

Top-1 Spearman Korelasyonu: 0.633

Top-5 Spearman Korelasyonu: 0.519

Sınıf 3:

Top-1 Pearson Korelasyonu: 0.614

Top-5 Pearson Korelasyonu: 0.581

Top-1 Spearman Korelasyonu: 0.596

Top-5 Spearman Korelasyonu: 0.505

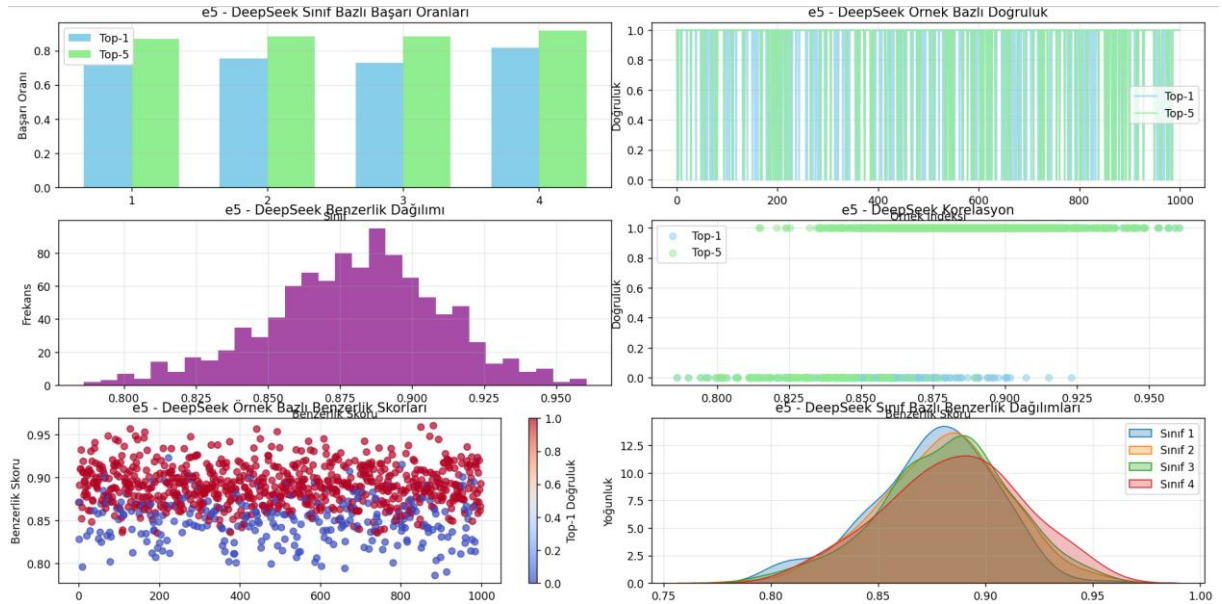
Sınıf 4:

Top-1 Pearson Korelasyonu: 0.647

Top-5 Pearson Korelasyonu: 0.555

Top-1 Spearman Korelasyonu: 0.615

Top-5 Spearman Korelasyonu: 0.464



Cosmos-Gpt4o

cosmos - GPT4O Korelasyon Sonuçları:

Top-1 Başarı Oranı: 77.70%

Top-5 Başarı Oranı: 94.10%

Top-1 Pearson Korelasyonu: 0.488

Top-5 Pearson Korelasyonu: 0.470

Top-1 Spearman Korelasyonu: 0.478

Top-5 Spearman Korelasyonu: 0.378

Örnek Bazlı Sonuçlar (İlk 10 örnek):

Örnek | Sınıf | Top-1 | Top-5 | Benzerlik Skoru

| | | | | |
|---|---|---|---|--------|
| 0 | 2 | 1 | 1 | 0.6995 |
| 1 | 1 | 0 | 0 | 0.5258 |
| 2 | 3 | 1 | 1 | 0.6714 |
| 3 | 3 | 0 | 0 | 0.5707 |
| 4 | 2 | 1 | 1 | 0.6607 |
| 5 | 3 | 1 | 1 | 0.8442 |
| 6 | 4 | 1 | 1 | 0.8173 |
| 7 | 1 | 1 | 1 | 0.7691 |
| 8 | 1 | 1 | 1 | 0.7018 |
| 9 | 3 | 1 | 1 | 0.6881 |

Sınıf Bazlı Başarı Oranları:

Sınıf 1:

Top-1 Başarı Oranı: 79.46%

Top-5 Başarı Oranı: 92.86%

Sınıf 2:

Top-1 Başarı Oranı: 77.45%

Top-5 Başarı Oranı: 94.47%

Sınıf 3:

Top-1 Başarı Oranı: 77.03%

Top-5 Başarı Oranı: 94.51%

Sınıf 4:

Top-1 Başarı Oranı: 77.55%

Top-5 Başarı Oranı: 93.88%

Sınıf Bazlı Korelasyonlar:

Sınıf 1:

Top-1 Pearson Korelasyonu: 0.526

Top-5 Pearson Korelasyonu: 0.517

Top-1 Spearman Korelasyonu: 0.506

Top-5 Spearman Korelasyonu: 0.417

Sınıf 2:

Top-1 Pearson Korelasyonu: 0.483

Top-5 Pearson Korelasyonu: 0.393

Top-1 Spearman Korelasyonu: 0.481

Top-5 Spearman Korelasyonu: 0.366

Sınıf 3:

Top-1 Pearson Korelasyonu: 0.477

Top-5 Pearson Korelasyonu: 0.475

Top-1 Spearman Korelasyonu: 0.468

Top-5 Spearman Korelasyonu: 0.369

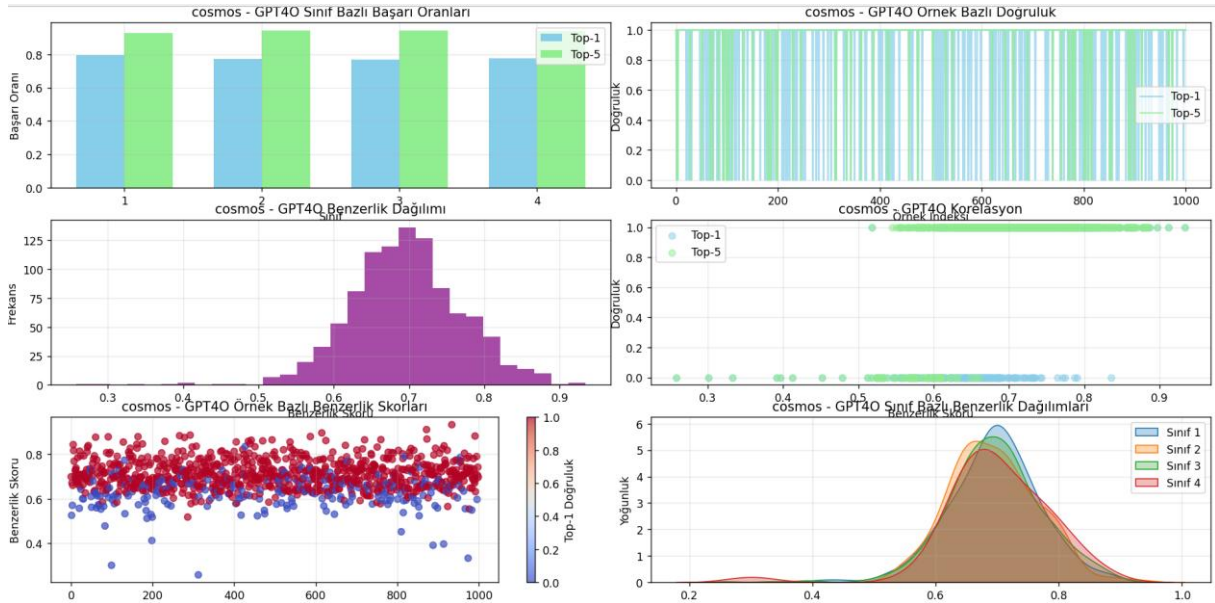
Sınıf 4:

Top-1 Pearson Korelasyonu: 0.485

Top-5 Pearson Korelasyonu: 0.554

Top-1 Spearman Korelasyonu: 0.450

Top-5 Spearman Korelasyonu: 0.349



Cosmos-Deepseek

cosmos - DeepSeek Korelasyon Sonuçları:

Top-1 Başarı Oranı: 72.40%

Top-5 Başarı Oranı: 88.70%

Top-1 Pearson Korelasyonu: 0.596

Top-5 Pearson Korelasyonu: 0.579

Top-1 Spearman Korelasyonu: 0.589

Top-5 Spearman Korelasyonu: 0.504

Örnek Bazlı Sonuçlar (İlk 10 örnek):

Örnek | Sınıf | Top-1 | Top-5 | Benzerlik Skoru

| | | | | |
|---|---|---|---|--------|
| 0 | 2 | 1 | 1 | 0.7113 |
| 1 | 1 | 0 | 0 | 0.5458 |
| 2 | 3 | 1 | 1 | 0.6627 |
| 3 | 3 | 0 | 1 | 0.6153 |
| 4 | 2 | 1 | 1 | 0.6970 |
| 5 | 3 | 1 | 1 | 0.7116 |
| 6 | 4 | 1 | 1 | 0.8187 |
| 7 | 1 | 1 | 1 | 0.5809 |
| 8 | 1 | 1 | 1 | 0.6416 |
| 9 | 3 | 0 | 0 | 0.5674 |

Sınıf Bazlı Başarı Oranları:

Sınıf 1:

Top-1 Başarı Oranı: 75.45%

Top-5 Başarı Oranı: 86.61%

Sınıf 2:

Top-1 Başarı Oranı: 68.94%

Top-5 Başarı Oranı: 85.53%

Sınıf 3:

Top-1 Başarı Oranı: 71.75%

Top-5 Başarı Oranı: 90.45%

Sınıf 4:

Top-1 Başarı Oranı: 81.63%

Top-5 Başarı Oranı: 95.92%

Sınıf Bazlı Korelasyonlar:

Sınıf 1:

Top-1 Pearson Korelasyonu: 0.629

Top-5 Pearson Korelasyonu: 0.681

Top-1 Spearman Korelasyonu: 0.586

Top-5 Spearman Korelasyonu: 0.553

Sınıf 2:

Top-1 Pearson Korelasyonu: 0.600

Top-5 Pearson Korelasyonu: 0.607

Top-1 Spearman Korelasyonu: 0.609

Top-5 Spearman Korelasyonu: 0.560

Sınıf 3:

Top-1 Pearson Korelasyonu: 0.582

Top-5 Pearson Korelasyonu: 0.512

Top-1 Spearman Korelasyonu: 0.581

Top-5 Spearman Korelasyonu: 0.458

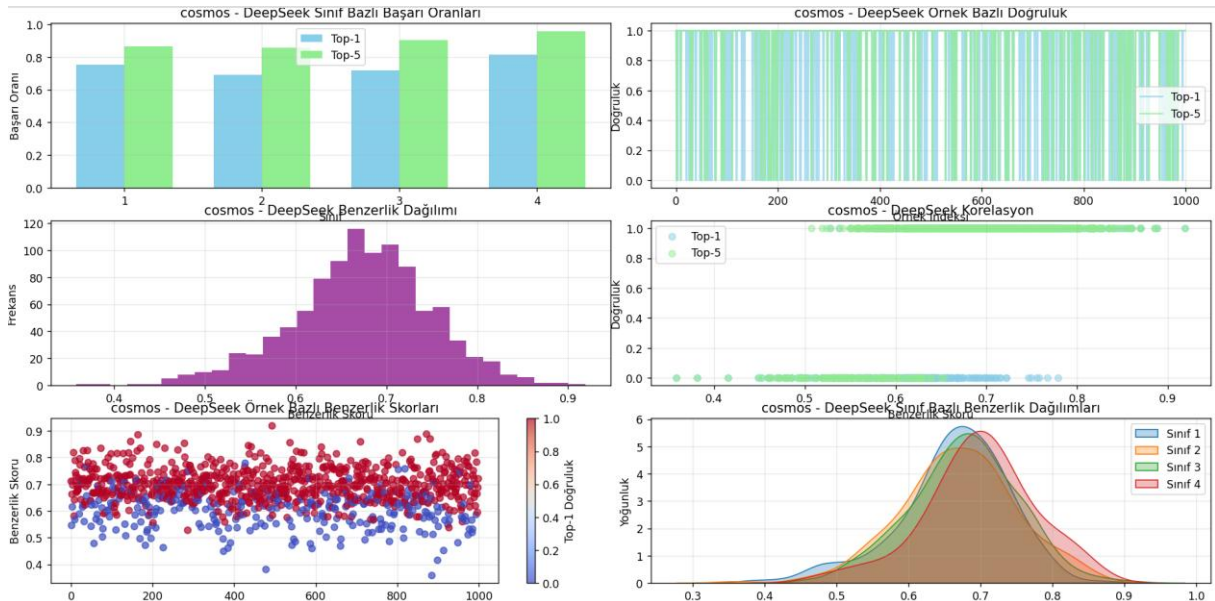
Sınıf 4:

Top-1 Pearson Korelasyonu: 0.646

Top-5 Pearson Korelasyonu: 0.516

Top-1 Spearman Korelasyonu: 0.607

Top-5 Spearman Korelasyonu: 0.343



Jina-Gpt4o

jina - GPT4O Korelasyon Sonuçları:

Top-1 Başarı Oranı: 89.50%

Top-5 Başarı Oranı: 97.10%

Top-1 Pearson Korelasyonu: 0.639

Top-5 Pearson Korelasyonu: 0.565

Top-1 Spearman Korelasyonu: 0.463

Top-5 Spearman Korelasyonu: 0.285

Örnek Bazlı Sonuçlar (İlk 10 örnek):

Örnek | Sınıf | Top-1 | Top-5 | Benzerlik Skoru

| | | | | |
|---|---|---|---|--------|
| 0 | 2 | 1 | 1 | 0.8882 |
| 1 | 1 | 1 | 1 | 0.7807 |
| 2 | 3 | 1 | 1 | 0.8167 |
| 3 | 3 | 1 | 1 | 0.8382 |
| 4 | 2 | 1 | 1 | 0.8884 |
| 5 | 3 | 1 | 1 | 0.9106 |
| 6 | 4 | 1 | 1 | 0.8402 |
| 7 | 1 | 1 | 1 | 0.8649 |
| 8 | 1 | 1 | 1 | 0.8056 |
| 9 | 3 | 1 | 1 | 0.8230 |

Sınıf Bazlı Başarı Oranları:

Sınıf 1:

Top-1 Başarı Oranı: 89.29%

Top-5 Başarı Oranı: 96.43%

Sınıf 2:

Top-1 Başarı Oranı: 88.09%

Top-5 Başarı Oranı: 98.72%

Sınıf 3:

Top-1 Başarı Oranı: 90.24%

Top-5 Başarı Oranı: 96.95%

Sınıf 4:

Top-1 Başarı Oranı: 89.80%

Top-5 Başarı Oranı: 93.88%

Sınıf Bazlı Korelasyonlar:

Sınıf 1:

Top-1 Pearson Korelasyonu: 0.703

Top-5 Pearson Korelasyonu: 0.609

Top-1 Spearman Korelasyonu: 0.493

Top-5 Spearman Korelasyonu: 0.316

Sınıf 2:

Top-1 Pearson Korelasyonu: 0.605

Top-5 Pearson Korelasyonu: 0.400

Top-1 Spearman Korelasyonu: 0.468

Top-5 Spearman Korelasyonu: 0.194

Sınıf 3:

Top-1 Pearson Korelasyonu: 0.620

Top-5 Pearson Korelasyonu: 0.574

Top-1 Spearman Korelasyonu: 0.442

Top-5 Spearman Korelasyonu: 0.290

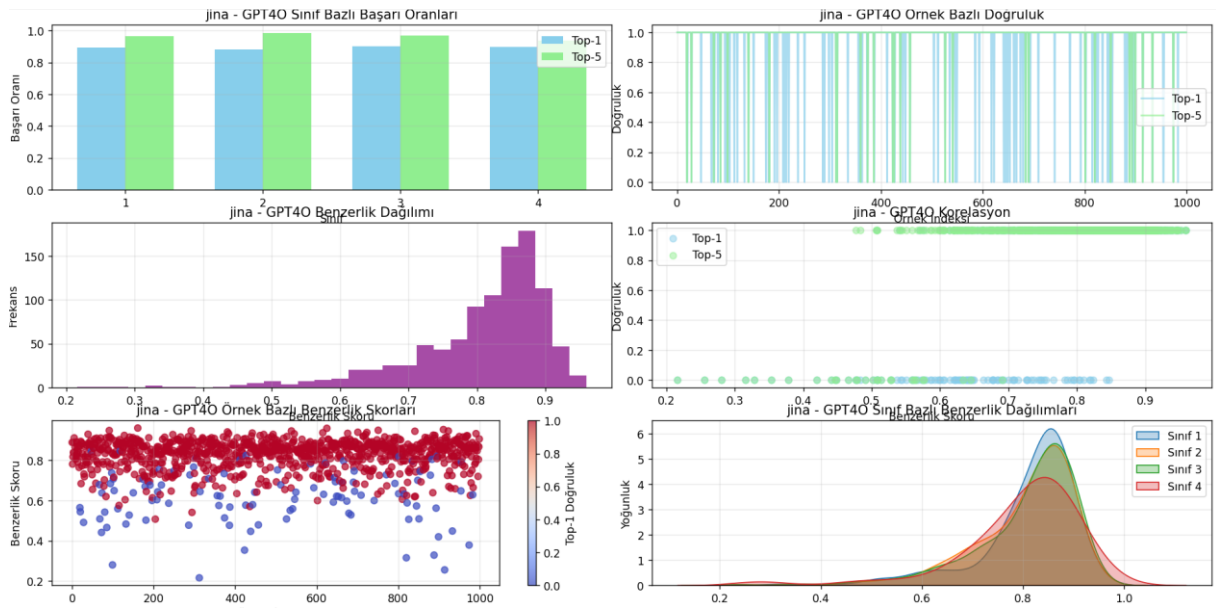
Sınıf 4:

Top-1 Pearson Korelasyonu: 0.766

Top-5 Pearson Korelasyonu: 0.771

Top-1 Spearman Korelasyonu: 0.520

Top-5 Spearman Korelasyonu: 0.415



Jina- Deepseek

jina - DeepSeek Korelasyon Sonuçları:

Top-1 Başarı Oranı: 84.50%

Top-5 Başarı Oranı: 95.50%

Top-1 Pearson Korelasyonu: 0.651

Top-5 Pearson Korelasyonu: 0.581

Top-1 Spearman Korelasyonu: 0.524

Top-5 Spearman Korelasyonu: 0.353

Örnek Bazlı Sonuçlar (İlk 10 örnek):

Örnek | Sınıf | Top-1 | Top-5 | Benzerlik Skoru

| | | | | |
|---|---|---|---|--------|
| 0 | 2 | 1 | 1 | 0.7895 |
| 1 | 1 | 1 | 1 | 0.8137 |
| 2 | 3 | 1 | 1 | 0.8349 |
| 3 | 3 | 1 | 1 | 0.8164 |
| 4 | 2 | 1 | 1 | 0.8829 |
| 5 | 3 | 1 | 1 | 0.8055 |
| 6 | 4 | 1 | 1 | 0.8240 |
| 7 | 1 | 1 | 1 | 0.7534 |
| 8 | 1 | 1 | 1 | 0.7897 |
| 9 | 3 | 0 | 0 | 0.5137 |

Sınıf Bazlı Başarı Oranları:

Sınıf 1:

Top-1 Başarı Oranı: 83.48%

Top-5 Başarı Oranı: 94.64%

Sınıf 2:

Top-1 Başarı Oranı: 84.26%

Top-5 Başarı Oranı: 97.02%

Sınıf 3:

Top-1 Başarı Oranı: 85.37%

Top-5 Başarı Oranı: 94.92%

Sınıf 4:

Top-1 Başarı Oranı: 81.63%

Top-5 Başarı Oranı: 97.96%

Sınıf Bazlı Korelasyonlar:

Sınıf 1:

Top-1 Pearson Korelasyonu: 0.679

Top-5 Pearson Korelasyonu: 0.611

Top-1 Spearman Korelasyonu: 0.534

Top-5 Spearman Korelasyonu: 0.378

Sınıf 2:

Top-1 Pearson Korelasyonu: 0.624

Top-5 Pearson Korelasyonu: 0.497

Top-1 Spearman Korelasyonu: 0.524

Top-5 Spearman Korelasyonu: 0.293

Sınıf 3:

Top-1 Pearson Korelasyonu: 0.652

Top-5 Pearson Korelasyonu: 0.609

Top-1 Spearman Korelasyonu: 0.510

Top-5 Spearman Korelasyonu: 0.375

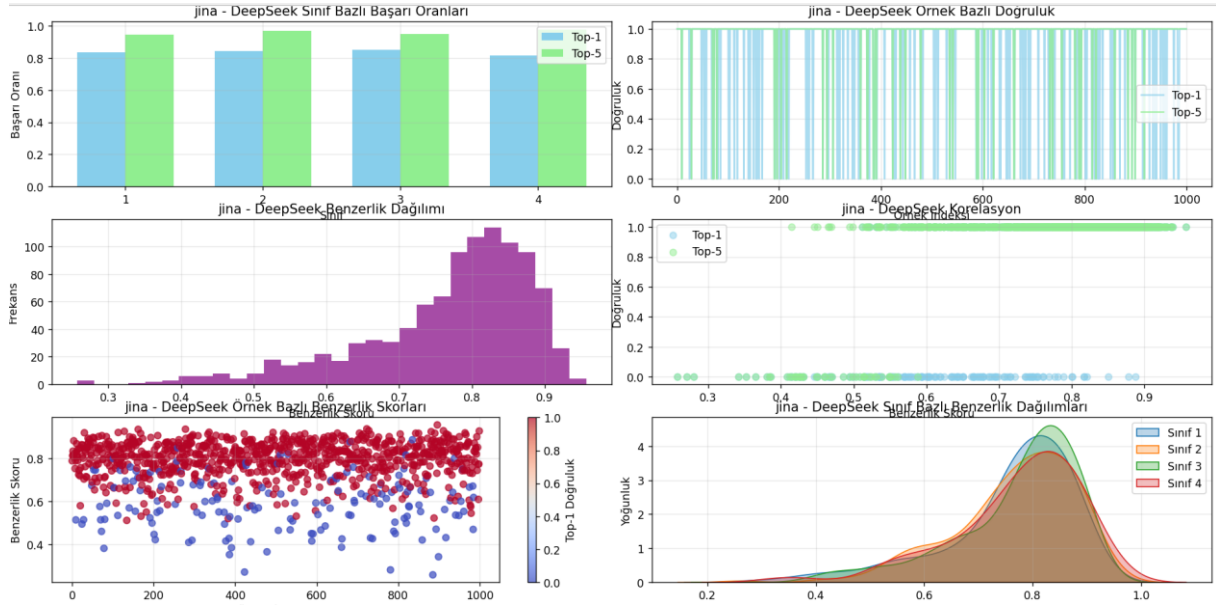
Sınıf 4:

Top-1 Pearson Korelasyonu: 0.640

Top-5 Pearson Korelasyonu: 0.518

Top-1 Spearman Korelasyonu: 0.566

Top-5 Spearman Korelasyonu: 0.245



Elde Edilen Sonuçların ve Grafiklerin Yorumlanması

1. Genel Performans Karşılaştırması:

- JINA Modeli en iyi performansı göstermiştir:
 - GPT4O ile: Top-1: %89.50, Top-5: %97.10
 - DeepSeek ile: Top-1: %84.50, Top-5: %95.50
- E5 Modeli orta düzeyde performansı göstermiştir:
 - GPT4O ile: Top-1: %77.60, Top-5: %93.60
 - DeepSeek ile: Top-1: %73.60, Top-5: %88.30
- COSMOS Modeli benzer performans göstermiştir:
 - GPT4O ile: Top-1: %77.70, Top-5: %94.10
 - DeepSeek ile: Top-1: %72.40, Top-5: %88.70

2. GPT4O vs DeepSeek Karşılaştırması:

- Tüm modellerde GPT4O ile daha yüksek başarı oranları elde edilmiştir.
- En büyük fark JINA modelinde görülmüştür (Top-1'de %5 fark)

3. Sınıf Bazlı Analiz:

- Sınıf 4 (ikisi de kötü) genellikle en düşük başarı oranlarına sahip olmuştur.
- Sınıf 1 (gpt4o daha iyi) ve Sınıf 2 (deepseek daha iyi) benzer performans göstermiştir.
- Sınıf 3 (ikisi de yeterince iyi) genellikle yüksek başarı oranlarına sahip olmuştur.

4. Korelasyon Analizi:

- Pearson korelasyonları genellikle Spearman'dan daha yüksek çıkmıştır.
- Top-1 korelasyonları Top-5'ten daha yüksek hesaplanmıştır.
- En yüksek korelasyonlar JINA modelinde görülmüştür.

5. Örnek Bazlı Analiz:

- Benzerlik skorları genellikle 0.5-0.9 arasında değişmiştir.
- JINA modelinde benzerlik skorları daha yüksek çıkmıştır.
- Yanlış tahminler genellikle düşük benzerlik skorlarına sahip olmuştur.

Bu sonuçlar, JINA modelinin diğer modellere göre daha iyi performans gösterdiğini ve GPT4O'nun DeepSeek'e göre daha iyi sonuçlar verdiğini göstermiştir. Ayrıca, Top-5 değerlendirmesinin Top-1'e göre daha güvenilir sonuçlar verdiği görülmüştür.

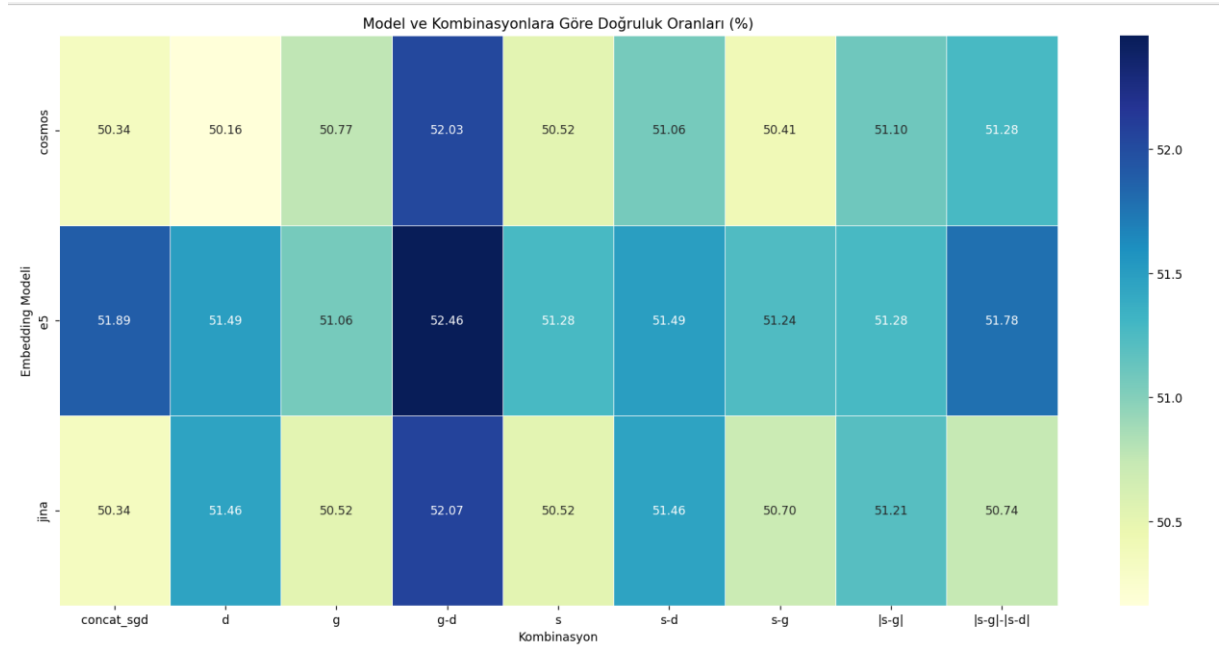
B-Hangisi İyi

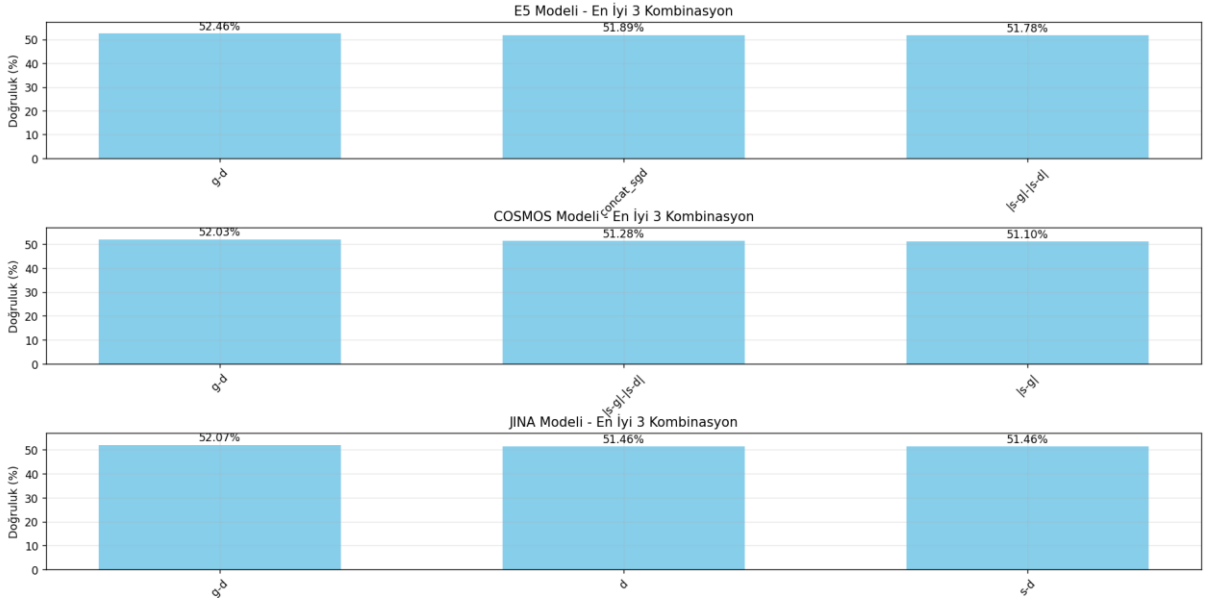
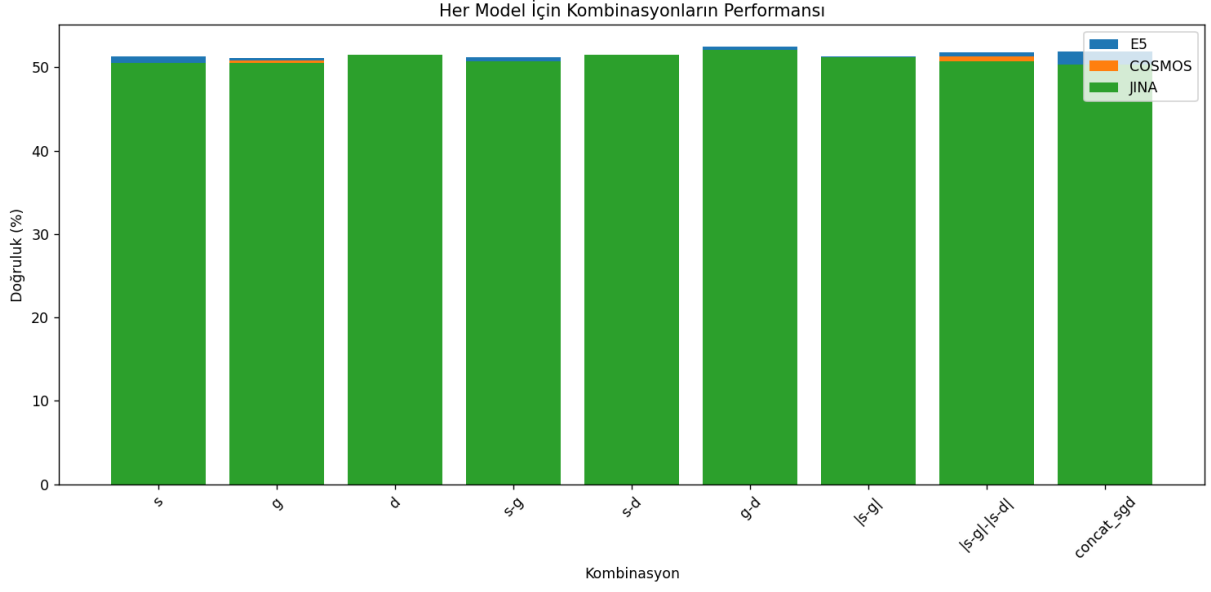
Bu soru için ilk önce tüm veri kümesi %80 eğitim %20 veri kümesi olarak 'EgitimTestAyirma.py' dosyası içerisinde ayrılmış ve 'train.xlsx', 'test.xlsx' şeklinde kaydedilmiştir. Daha sonra 'embeddingTumVeri.py' dosyası içerisinde eğitim ve test verilerinin her model için ve her modelde soru, gpt4o, deepseek featureları için embeddingleri çıkarılmıştır. Çıkarılan embeddingler 'embeddingsTumveri' klasoru içerisine kaydedilmiştir. 'b_hangisi_ iyi.py' içerisinde öncelikle, eğitim ve test setleri yüklenmiştir. Ardından, 9 farklı kombinasyon (örneğin, s-g, s-d, |s-g| vb.) oluşturulmuştur. Her bir model ve kombinasyon için Logistic Regression algoritması eğitilmiştir. Eğitilen modellerin performansı değerlendirilmiştir.

Görselleştirme kısmında ise ısı haritası (heatmap) ve çubuk grafikler (bar plots) kullanılarak model performansları görsel olarak analiz edilmiştir. En iyi 3 kombinasyon detaylı şekilde incelenmiştir.

Sonuç olarak, her bir model ve kombinasyon için doğruluk oranları hesaplanmış ve en iyi performansı gösteren kombinasyonlar belirlenmiştir. Detaylı görsel analizlerle hangi yaklaşımın daha iyi sonuç verdiği ortaya konmuştur.

Kısacası, bu kod farklı embedding modelleri ve kombinasyonlarının performansını karşılaştırmalı bir şekilde analiz etmekte ve hangi yaklaşımın daha başarılı olduğunu göstermiştir.





Sonuçların Yorumlanması

1. Genel Performans:

- Tüm modellerde doğruluk oranları %50 ile %52.5 arasında değişmiştir.
- En yüksek doğruluk oranı E5 modelinde %52.46 ile elde edilmiştir.
- En düşük doğruluk oranı ise COSMOS modelinde %50.16 ile elde edilmiştir.

2. En İyi Kombinasyonlar:

E5 Modeli:

- En iyi kombinasyon: g-d (%52.46)

2. İkinci en iyi kombinasyon: concat_sgd (%51.89)
3. Üçüncü en iyi kombinasyon: |s-g|-|s-d| (%51.78)

COSMOS Modeli:

1. En iyi kombinasyon: g-d (%52.03)
2. İkinci en iyi kombinasyon: |s-g|-|s-d| (%51.28)
3. Üçüncü en iyi kombinasyon: |s-g| (%51.10)

JINA Modeli:

1. En iyi kombinasyon: g-d (%52.07)
2. İkinci en iyi kombinasyon: d ve s-d (%51.46)
3. Üçüncü en iyi kombinasyon: |s-g| (%51.21)

3. Sonuç Olarak:

- g-d (GPT4O ve DeepSeek embeddinglerinin farkı) kombinasyonu tüm modellerde en iyi sonucu vermiştir.
- Tek embedding kullanımı (s, g, d) genellikle daha düşük performans göstermiştir.
- Fark kombinasyonları (s-g, s-d, g-d) genellikle daha iyi sonuçlar sağlamıştır.
- Mutlak fark kombinasyonları (|s-g|, |s-g|-|s-d|) orta düzeyde performans sergilemiştir.
- E5 Modeli: En yüksek maksimum doğruluk oranı (%52.46) ile en iyi performansı gösteren modeldir.
- JINA Modeli: En tutarlı performansı sunmuş, çoğu kombinasyonda doğruluk oranları %50 ile %52 arasında değişmiştir.
- COSMOS Modeli: En düşük minimum doğruluk oranına sahip olmuş, (%50.16) bu da modelin daha az tutarlı olduğunu göstermektedir.
- Tüm doğruluk oranları %50 ile %52.5 arasında kalmış, bu da modelin tahminlerinin rastgele tahminden çok daha iyi olmadığını göstermektedir.
- Sınıflandırma performansı genel olarak düşüktür, bu da mevcut modelin daha güçlü bir yaklaşım gerektirdiğini işaret etmektedir.
- Daha gelişmiş modeller veya farklı yaklaşımlar denenmesi gerektiği sonucu ortaya çıkmıştır.