



**T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLER ENSTİTÜSÜ
BİLGİSAYAR MÜHENDİSLİĞİ
TEZLİ YÜKSEK LİSANS PROGRAMI**

BLM 5109: KOLLEKTİF ÖĞRENME

Retrieval Ensemble

**AHMED UĞUR
23501027**

**METİN USLU
235B7014**

**Öğretim Üyesi
Prof. Dr. Mehmet Fatih AMASYALI**

**İSTANBUL
Aralık, 2024**

Retrieval Ensemble

Özet

Bu çalışmada, Türkçe dilinde retrieval ensemble yöntemleri kullanılarak sorulara doğru cevapların seçilmesi hedeflenmiştir. Çalışmada, 51600 adet Türkçe talimat ve cevap çiftinden oluşan bir veri kümesi üzerinde deneyler gerçekleştirilmiş ve 2000 örnek üzerinde analizler yapılmıştır. Sorular ve cevaplar arasındaki benzerliklerin belirlenmesi için 5 farklı embedding yöntemi kullanılmış, bu yöntemlerin kararları 3 farklı ensemble yöntemiyle birleştirilmiştir: Majority Voting, Average Cosine Similarity ve Weighted Cosine Similarity.

Deneysel sonuçlar, ensemble yöntemlerinin tekil modellere kıyasla daha yüksek doğruluk oranlarına ulaştığını göstermiştir. Weighted Cosine Similarity yöntemi, en yüksek Top-1 doğruluk oranına ulaşırken, Majority Voting yöntemi Top-5 doğruluk oranında en iyi performansı göstermiştir. Çalışma, retrieval sistemlerinde ensemble yöntemlerinin önemli bir katkı sunduğunu ve Türkçe dilinde özgün bir çerçeve sunduğunu ortaya koymuştur. Bu bulgular, retrieval tabanlı sistemlerde daha karmaşık yöntemlerin geliştirilmesi ve farklı veri kümeleri ve ensemble yöntemleri için bir temel sağlamaktadır.

1. Giriş

Bilgiye erişim hızının giderek arttığı günümüzde, büyük veri setlerinden doğru ve ilgili bilgiyi hızlı bir şekilde elde etmek, hem akademik hem de endüstriyel uygulamalar için kritik bir öneme sahiptir. Bu bağlamda, bilgi erişim sistemleri ve soru-cevap modelleri, verimli bilgi çıkarımı için önemli araçlar haline gelmiştir. Ancak, büyük veri setleri üzerinde gerçekleştirilen bilgi geri çağırma (retrieval) işlemlerinde tek bir model veya yöntem genellikle optimal sonuçlar sunmamaktadır. Bu durum, farklı yöntemlerin birleştirilerek daha güçlü ve genel bir çözüm sunmayı amaçlayan ensemble yaklaşımlarının kullanımını teşvik etmiştir.

Bu çalışmada, retrieval ensemble yöntemleri kullanılarak bir sorunun doğru cevabını bir aday havuzundan bulma problemi ele alınmıştır. Çalışmada, 2000 farklı cevap adayı içeren bir veri seti üzerinde, farklı embedding ve ensemble yöntemlerinin tekil ve birleştirilmiş performansları incelenmiştir. Amaç, tekil modellerin doğruluğunu artırmak ve karar verme sürecini daha güvenilir hale getirmektir.

Çalışmada, top-1 ve top-5 performans ölçütlerine dayalı olarak yöntemlerin etkinliğini değerlendirmiştir. Top-1, en yüksek olasılıkla tahmin edilen cevabın doğru olup olmadığını ölçerken, Top-5, en doğru cevabın ilk beş aday arasında yer alıp almadığını değerlendirmektedir. Bu ölçütler, sistemlerin doğruluk ve esneklik açısından performansını analiz etmek için güçlü bir çerçeve sunar.

Bu bağlamda, üç farklı ensemble yöntemi uygulanmıştır: Majority Voting (Çoğunluk Oyu), Average Cosine Similarity (Ortalama Kosinüs Benzerliği) ve Weighted Cosine Similarity (Kosinüs Benzerliğinin Ağırlıklı Ortalaması). Çalışmada, her bir yöntemin detaylı analizleri yapılarak, farklı embedding teknikleri ve bu tekniklerin kombinasyonlarının nasıl performans gösterdiği araştırılmıştır. Çalışmanın çıktıları, retrieval süreçlerinde ensemble yaklaşımlarının güçlü bir alternatif olabileceğini göstermeyi amaçlamaktadır.

2. Veri Kümesi

Bu çalışmada kullanılan veri kümesi, **Hugging Face** platformunda sağlanan **Turkish Instruction Dataset** adlı bir veri setidir. Veri seti, toplamda 51.600 satırdan oluşan ve her bir satırında "talimat", "giriş" ve "çıkıtı" sütunlarını içeren bir Türkçe metin veri kümesidir. Bu sütunlar sırasıyla bir görevi tanımlayan talimatı, isteğe bağlı olarak verilen giriş bilgisini ve bu bilgilere göre oluşturulan çıktıyı temsil etmektedir. Veri kümesi, dil işleme görevleri için zengin bir kaynak sağlamakta ve çeşitli metin türleri ile kapsamlı bir analiz imkânı sunmaktadır.

Bu çalışmanın amaçlarına uygun olarak veri setine bazı ön işlemler uygulanmıştır. Öncelikle veri集中的 sütun isimleri Türkçe karakter içermeyecek şekilde "talimat, giriş, cevap" şeklinde düzenlenmiştir. Her satırda isteğe bağlı olarak verilen giriş verisi bulunmadığı için dağılımı giriş bilgisine yönelik olarak korumak amacıyla 1000 satır giriş verisi içeren ve 1000 satır giriş verisi içermeyen bir şekilde alt bir veri kümesi ile sınırlandırılmıştır.

Tablo 1 – Türkçe Talimatlar Veri Kümesi

talimat	giris	cikti
Bileşik bir cümle oluşturun.	Sonuçlar konusunda endişeliydi; çok çalışmıştı.	Çok çalışmıştı ve sonuç olarak sonuçlar konusunda endişeliydi.
Verilen cümleyi uygun dilbilgisine sahip olacak şekilde yeniden düzenleyin.	Köpek takip ediliyordu.	Köpek onu takip ediyordu.
Ünlü bir İtalyan şarkıcı söyleyin.	NaN	Ünlü bir İtalyan şarkıcı Andrea Bocelli'dir.
Verilen problem için daha yaratıcı fikirlerin bir listesini oluşturun.	Çevrimiçi satışlar nasıl artırılır?	1. Devam eden müşterilere indirimler sunmak 2. Kullanıcı dostu bir web sitesi tasarlamak 3. Görünürlüğü artırmak için etkileyicilerden yararlanmak 4. Potansiyel müşterileri çekmek ve etkileşimde bulunmak için etkileşimli içerik oluşturmak 5. Alışveriş sepetini terk eden müşteriler için yeniden hedefleme kampanyaları kullanmak 6. Ücretsiz kargo sunmak veya diğer teşvikler...
Aşağıdaki metni inceleyin ve herhangi bir intihal belirtisi olup olmadığını tespit edin.	İnternet, görünüşte sonsuz sayıda web sitesi, blog ve kaynakla geniş bir bilgi kaynağı haline geldi. Çoğu durumda, bu web sitelerindeki içeriğin bir başkası tarafından oluşturulduğu ve sahiplenildiği gerçeğini gözden kaçırmak kolaydır.	Böyle bir önyargı tespit edilmedi.

Talimat ve giriş sütunları birleştirilerek “soru” isminde yeni bir sütun oluşturulmuştur. Bu sütun retrieval görevinde kullanılacak soru metinlerini temsil etmektedir. Tablo 1’de veri kümesinden rastgele 5 örnek gösterilmektedir. Veri kümesinin retrieval görevinde sorular; talimat ve giriş sütunlarının birleştirilmesiyle elde edilen “soru” sütunundaki metinler, cevaplar ise “cikti” sütununda yer alan metinler kullanılarak yapılandırılmıştır.

Veri kümesinin retrieval probleminde kullanım amacı, verilen bir soruya karşılık gelen doğru cevabın, 2000 cevap adayı arasından en iyi şekilde bulunmasını sağlamaktır. Bu doğrultuda, veri setindeki her bir soru ve doğru cevabı benzersiz bir şekilde eşleşmektedir.

3. Deneysel Analiz

Bu çalışmada, retrieval görevinde kullanılan farklı embedding yöntemlerinin ve bu yöntemlerin ensemble stratejileriyle birleştirilmesinin etkinliği değerlendirilmiştir. Sorular ve cevaplar, beş farklı embedding modeli kullanılarak vektörlere dönüştürülmüştür. Bu süreçte her bir model için sorular ve cevaplar ayrı ayrı işlenmiş ve yüksek boyutlu vektör uzaylarında temsil edilmiştir. Embedding yöntemleri, metinlerin sayısal birer vektör olarak temsil edilmesini sağlayarak retrieval görevinde benzerlik ölçütlerinin uygulanmasını mümkün kılar.

Tablo 2’de kullanılan temsil yöntemlerine ait kullanılan parametre sayıları, vektör boyutları ve destekledikleri dillere ait bilgiler verilmiştir.

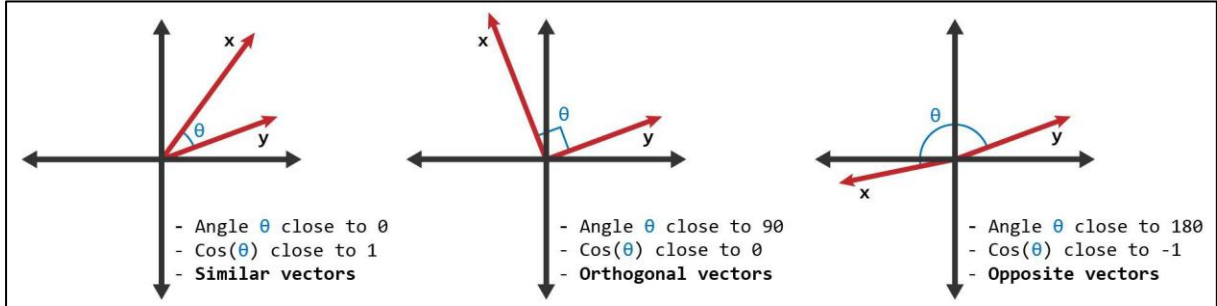
Tablo 2 – Temsil Yöntemlerine Ait Bilgiler

Model Adı	Parametre Sayısı	Vektör Boyutu	Desteklenen Diller
MiniLM-L12-v2	33.4M	384	İngilizce
Jina-v3	572M	1536	Çok Dilli (50+ dil)
E5-Large	560M	1024	Çok Dilli (50+ dil)
BGE-M3	-	1024	İngilizce, Çince
Turkish-ColBERT	111M	128	Türkçe
GTE-Large	335M	768	İngilizce

Soruların ve cevapların vektör temsilleri arasında benzerlik ölçümü yapmak için kosinüs benzerliği (cosine similarity) metriği kullanılmıştır. Her bir soru için, 2000 cevap adayı arasında doğru cevaba en yakın olan vektörler sıralanmıştır.

Kosinüs benzerliği, iki vektör arasındaki benzerliği ölçmek için kullanılır. Bu hesaplama iki vektör arasındaki açıya dayalı olarak hesaplanır. Aralarındaki açı küçüldükçe vektörlerin benzer olduğu kabul edilir. Şekil 1’de kosinüs benzerliğinin üç temel durumu gösterilmiştir. İlk durumda, iki vektör arasındaki açı (θ) küçük olduğunda (0° ’a yakın) kosinüs değeri 1’e yaklaşır. Bu durum, vektörlerin birbirine paralel ve aynı yönde olduğu ifade eder. İkinci durumda, açı (θ) yaklaşık 90° derece olduğunda kosinüs değeri 0’a yakın olur. Bu durum vektörlerin birbirine dik olduğunu ve aralarında ilişki olmadığını gösterir. Üçüncü durumda ise açı (θ) yaklaşık 180° derece olduğunda kosinüs değeri -1’e yaklaşır. Bu durum da vektörlerin birbirine tamamen zıt yönlerde olduğunu ifade eder.

Şekil 1 – Kosinüs Benzerliği Üç Temel Durumu



Yapılan analizde tekil modellerin performanslarını ölçmek için top-1 ve top-5 accuracy kullanılmıştır. Çalışma kapsamında her bir soru için oluşan cevap tahminleri kosinüs benzerliklerine göre sıralandığında ilk sıradaki cevap, gerçek cevap ise top-1 başarısı bir bir artar. Eğer doğru cevap ilk beş sırada yer alıyorsa ise top-5 başarısı bir artırılmıştır.

Tablo 3’te retrieval görevinde kullanılmak üzere veri kümesinde ön işlemler uygulanan örnek bir soru-cevap çifti yer almaktadır. Bu örnek soru özelinde tekil modellerin tahmini cevapları ise top-5 listesi olarak Tablo 4’te gösterilmiştir.

Tablo 3 – Türkçe Talimatlar Veri Kümesi Örnek Soru-Cevap

Soru	Cevap
Belirli bir konu hakkında bir makale yazın. Konu: Zaman Yönetimi	Zaman yönetimi, hayatta başarılı olmak için sahip olunması gereken önemli bir beceridir. Gününüzden en iyi şekilde yararlanabilmeniz için görevlere öncelik vermenizi ve zamanınızı en iyi şekilde değerlendirmenizi sağlar...

Tablo 4 – Örnek Soru için Tekil Modellerin Çıktıları

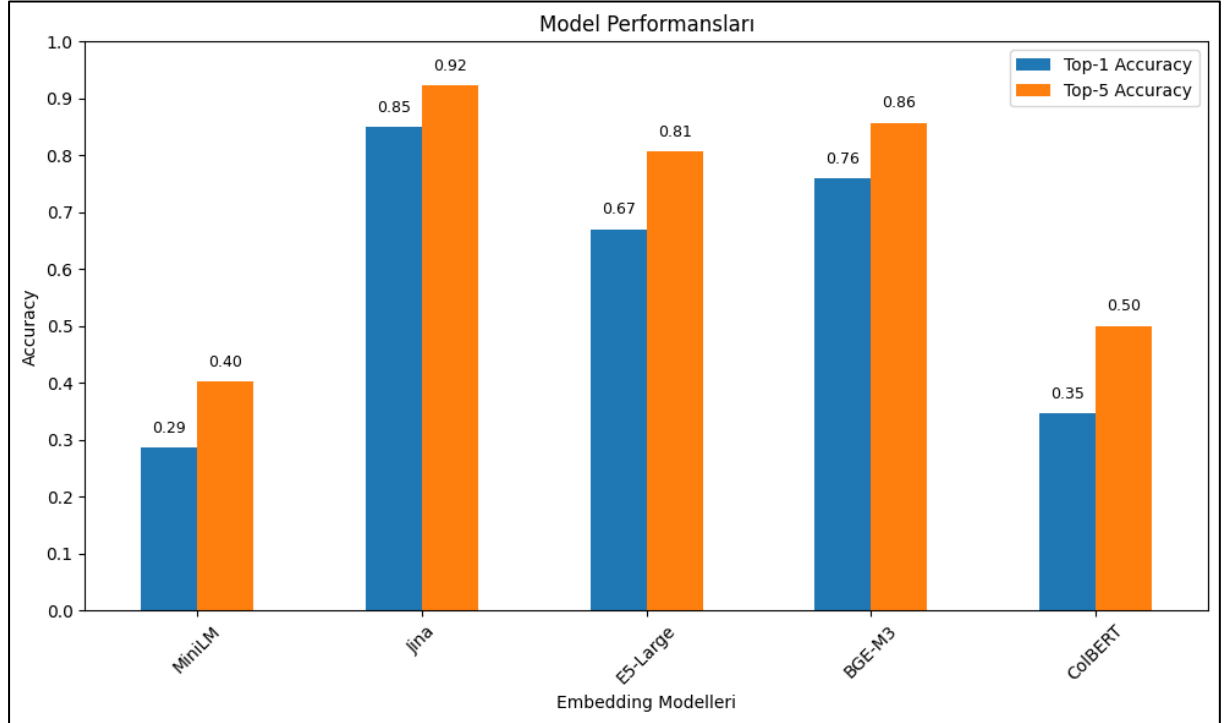
Retrieval Çıktıları		
Model Adı	Kosinüs Benzerliği	Tahmini Cevaplar
MiniLM-L12 v2	0.7902	Makaleleri bulmak ve yönetmek için kullanılan bir kütüphane yazılımı...
	0.7861	Öğretmen köpek hakkında bir makale yazdı.
	0.7718	Belirli bir durumdaki kişi muhtemelen bir kaygı bozukluğu yaşıyor.
	0.6929	Yapay zeka, bir makinenin akıllı davranışı taklit etme yeteneğidir...
	0.6924	Yarın üretken olabilmem için dinlenmeye ve şarj olmaya biraz zaman...
Jina-v3	0.7149	Zaman yönetimi, hayatta başarılı olmak için sahip olunması gereken önemli bir beceridir. Gününüzden en iyi şekilde yararlanabilmeniz için görevlere öncelik vermenizi ve zamanınızı en iyi şekilde değerlendirmenizi sağlar...
	0.4552	Zamanlar zor olabilir, ancak iyi seçilmiş teknikler yardımcı olur.
	0.4308	Bir algoritmanın zaman karmaşıklığını, kodu inceleyerek ve girdinin...
	0.4120	Kısıtlı sürede bir projeyi tamamlama sorununu çözmek için düşünülebilecek üç çözüm vardır...
	0.4112	"Vakit nakittir" atasözü, kişinin zamanının değerli bir varlık olduğunu...
E5-Large	0.8708	Zaman yönetimi, hayatta başarılı olmak için sahip olunması gereken önemli bir beceridir. Gününüzden en iyi şekilde yararlanabilmeniz için görevlere öncelik vermenizi ve zamanınızı en iyi şekilde değerlendirmenizi sağlar...
	0.8516	Metin oluşturma görevi.
	0.8516	Zamanda Yolculuk: Bir Muhasebecinin İnanılmaz Macerası.
	0.8498	Verilen metnin ana teması barıştır.
	0.8421	Model, verilen metni "Aile" konusuna ait olarak sınıflandırmalıdır.
BGE-M3	0.7137	Zaman yönetimi, hayatta başarılı olmak için sahip olunması gereken önemli bir beceridir. Gününüzden en iyi şekilde yararlanabilmeniz için görevlere öncelik vermenizi ve zamanınızı en iyi şekilde değerlendirmenizi sağlar...
	0.5550	Kişisel Yazı.
	0.5377	Eğitim ve aile yükümlülüklerini dengelemenin bir yolu, bir program oluşturmak ve önceden plan yapmaktır...
	0.5354	Kısıtlı sürede bir projeyi tamamlama sorununu çözmek için...
	0.5320	Herkese günaydın. Bugün kişisel farkındalığın öneminden bahsetmek...
Turkish-CoBERT	0.8741	Güzel yazı.
	0.8325	Model, verilen metni "Aile" konusuna ait olarak sınıflandırmalıdır.
	0.8279	1. Coursera: IBM tarafından "Yapay Zekaya Giriş" 2. Udacity: "Yapay Zekaya Giriş" kursu 3. MIT OpenCourseware: "Yapay Zekaya Giriş" 4. edX: Harvard ve IBM tarafından "Yapay Zekaya Giriş"...
	0.8198	SEÇİN * NEREDE konu = 'bahçıvanlık' makalelerinden alaka düzeyine göre SİPARİŞ TANININ;
	0.8173	Tavsiye edeceğim psikoloji üzerine üç kitap: 1) Daniel Kahneman'ın Hızlı ve Yavaş Düşünme; 2) Sannd'in Sonu, Dr. Paul Dolan; ve 3) Alışkanlığın Gücü, Charles Duhigg.

Tablo 5 – Tekil Modellerin Başarı ve Hataları

Model Name	Top-1 Accuracy	Top-5 Accuracy	Top-1 Errors/Rate	Top-5 Errors/Rate
MiniLM-L12 v2	0.2865	0.4025	1427 (71.35%)	1195 (59.75%)
Jina-v3	0.8495	0.9235	301 (15.05%)	153 (7.65%)
E5-Large	0.6705	0.8075	659 (32.95%)	385 (19.25%)
BGE-M3	0.7600	0.8575	480 (24.00%)	285 (14.25%)
Turkish-ColBERT	0.3470	0.5005	1306 (65.30%)	999 (49.95%)

Tablo 5’te beş farklı embedding yöntemiyle oluşturulan modellerin veri setindeki tüm sorulara verdikleri cevaplar üzerinden ölçülen performans metrikleri yer almaktadır. Ayrıca bu metrikler grafik olarak da Şekil 2’de yer almaktadır. Buna göre top-1 ve top-5 ölçümünde en yüksek başarıyı Jina-v3 modeli oluşturmuştur. Ardından başarıyı BGE-M3 ve E5-Large modelleri takip etmektedir. MiniLM-L12 v2 modeli ise tekil modeller arasında en az başarıyı oluşturmuştur. Bunun nedeni, modelin daha az parametre sayısı olması, sadece İngilizce dil desteği olması ve amacının kısa paragraflar veya tek bir cümle kodlamak olması gibi nedenlerle açıklanabilir.

Şekil 2 – Tekil Model Performansları



Tekil modellerin kararlarını birleştirmek ve genel performansı arttırmak için ensemble yöntemleri uygulanmıştır. Çalışmada üç farklı ensemble stratejisi izlenmiştir olup bu yöntemlerin tekil modellerin performansına kıyasla getirdiği avantajlar analiz edilmiştir.

Majority Voting

Bu yöntemde her bir örnek soru için tekil modelin top-1 ve top-5 cevap aday listelerine bakılır. Her bir listedeki aday diğer modellerin de listelerinde yer alıyorsa oyu 1 artırılır. En yüksek oya sahip cevap aday ensemble sonucu olarak belirlenir.

$$\hat{y} = \operatorname{argmax}_i \left(\sum_{j=1}^M 1 [i \in \text{Top-}X_j] \right)$$

$1[i \in \text{Top-}X_j]$: i'inci cevabın j'inci modelin Top-X tahmininde olup olmadığını kontrol eden bir gösterge fonksiyonudur. (1 ise var, 0 ise yoktur.)

Average Cosine Similarity

Bu yöntemde her bir örnek soru için tekil modellerin cevap adaylarına ilişkin hesaplanan kosinüs benzerliklerinin ortalaması alınır. En yüksek ortalamaaya ait ilk cevap top-1, ilk beş cevap ise top-5 ensemble sonucu olarak belirlenir.

$$\hat{y} = \operatorname{argmax}_i \left(\frac{1}{M} \sum_{j=1}^M S_{ij} \right)$$

S_{ij} : j-inci model tarafından i-inci cevap için hesaplanan kosinüs benzerliği skorunu ifade eder.

Weighted Cosine Similarity

Bu yöntemde ise her bir tekil modelin top-1 doğruluk oranları normalize edilerek toplam 5 farklı modele ait elde edilen değerlerin toplamı 1 olacak şekilde w_j olarak ağırlıklandırılmıştır. Tekil modellerin kararları birleştirilirken başarısı daha yüksek olan modelin ağırlığı daha yüksek olduğu için karara olan etkisi daha fazla olacaktır.

Tüm modellerin kosinüs benzerlikleri kendi ağırlığı ile çarpıldığında en yüksek skora sahip ilk cevap top-1, ilk beş cevap ise top-5 ensemble sonucu olarak belirlenir.

$$\hat{y} = \operatorname{argmax}_i \left(\sum_{j=1}^M w_j \cdot S_{ij} \right)$$

w_j : j-inci modelin ağırlığını ifade etmektedir. j-inci modelin tekil top-1 doğruluğunun normalize edilmiş halidir.

S_{ij} : j-inci model tarafından i-inci cevap için hesaplanan kosinüs benzerliği skorudur.

Tablo 6 – Tekil Modellerin Başarı ve Hataları

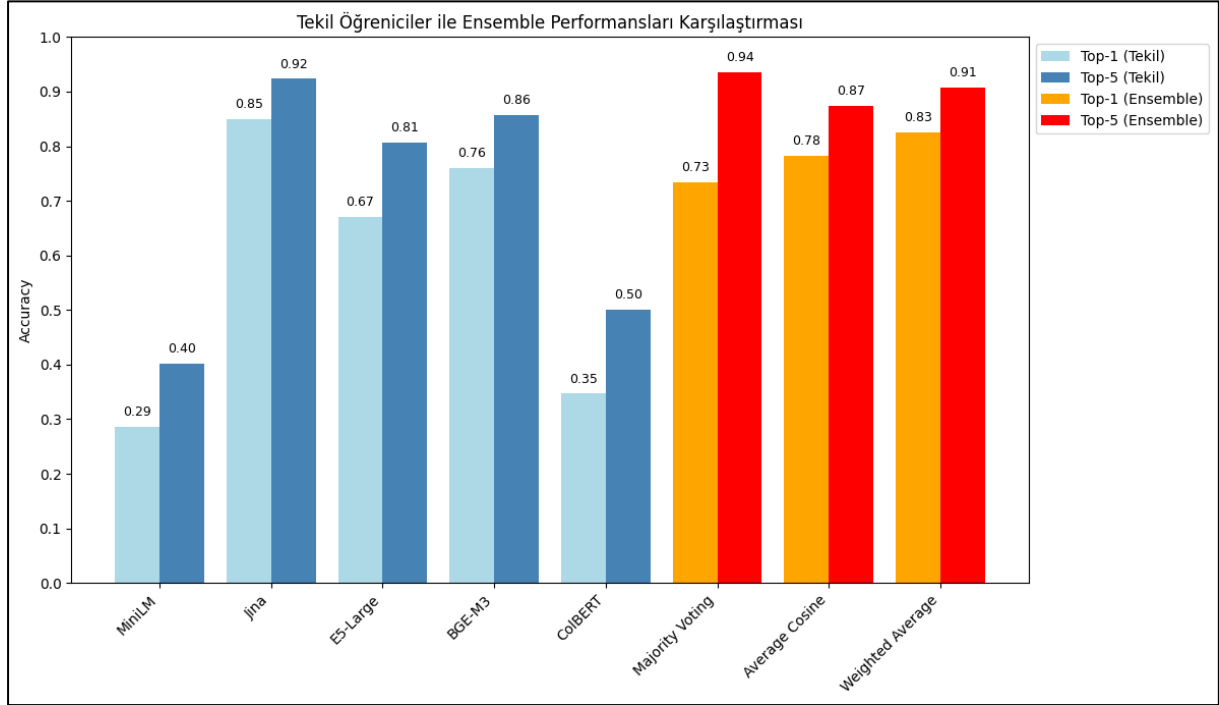
Ensemble Method Name	Top-1 Accuracy	Top-5 Accuracy
Majority Voting	0.7340	0.9355
Average Cosine Similarity	0.7830	0.8740
Weighted Cosine Similarity	0.8260	0.9075

Tablo 6’da uygulanan üç farklı ensemble stratejisine ait top-1 ve top-5 accuracy sonuçları yer almaktadır. Top-1 başarı sıralamasında Weighted Cosine Similarity yöntemi en iyi performansı sergilemektedir. Bu yöntem tekil modellere bireysel başarılarına göre ağırlık belirlenerek hesaplama yapıldığı için daha dengeli ve tutarlı bir şekilde genelleme yapabildiği anlamına gelmektedir. Ancak, yöntemin uyguladığı ağırlıklandırma top-5 başarı ölçümünde oylama yerine doğrudan kosinüs benzerlik skorlarına dayalı bir sıralama yaptığı için top-5 listesindeki çeşitliliği azaltabilir ve bazı doğru cevapların liste dışı kalmasına neden olabilir.

Majority Voting yöntemi ise top-5 başarı sıralamasında en iyi performansı sergilemektedir. Oylama tabanlı olduğu için düşük performanslı modellerin etkisi burada avantaj olabilmektedir. Çoğunluk oyu, birden fazla modelden gelen hatalı tahminleri dengeleyerek top-5 sıralamasında doğru cevabın yer almasını sağlar. Bir cevabın, top-5 adaylar içinde yer almasının yeterli kabul edilebileceği görevlerde oldukça etkili bir yöntem olarak kullanılabilir.

Şekil 3'te uygulanan üç farklı ensemble stratejisi ile tekil modellerin top-1 ve top-5 başarı sonuçları grafik üzerinden karşılaştırılmıştır.

Şekil 3 – Ensemble Yöntemler ile Tekil Modellerin Performans Karşılaştırması



Tablo 7 – Örnek Soru için Ensemble Yöntemlerine Ait Çıktılar

Ensemble Sonuçları		
Yöntem Adı	Skor	Tahmini Cevap
Majority Voting	3	Zaman yönetimi, hayatta başarılı olmak için sahip olunması gereken önemli bir beceridir. Gününüzden en iyi şekilde yararlanabilmeniz için görevlere öncelik vermenizi ve zamanınızı en iyi şekilde değerlendirmenizi sağlar...
	2	Kısıtlı sürede bir projeyi tamamlama sorununu çözmek için düşünülebilecek üç çözüm vardır. İlk çözüm, zamanı dikkatli bir şekilde planlamak...
	2	Model, verilen metni “Aile” konusuna ait olarak sınıflandırmalıdır.
	1	Makaleleri bulmak ve yönetmek için kullanılan bir kütüphane yazılımı EndNote'tur.
	1	Öğretmen köpek hakkında bir makale yazdı.
Average Cosine Similarity	0.7004	Zaman yönetimi, hayatta başarılı olmak için sahip olunması gereken önemli bir beceridir. Gününüzden en iyi şekilde yararlanabilmeniz için görevlere öncelik vermenizi ve zamanınızı en iyi şekilde değerlendirmenizi sağlar...
	0.6382	Zamanda Yolculuk: Bir Muhasebecinin İnanılmaz Macerası.
	0.6235	Kişisel Yazı.
	0.6208	Öğretmen köpek hakkında bir makale yazdı.
	0.6152	Makaleleri bulmak ve yönetmek için kullanılan bir kütüphane yazılımı EndNote'tur.

Weighted Cosine Similarity	0.7250	Zaman yönetimi, hayatta başarılı olmak için sahip olunması gereken önemli bir beceridir. Gününüzden en iyi şekilde yararlanabilmeniz için görevlere öncelik vermenizi ve zamanınızı en iyi şekilde değerlendirmenizi sağlar...
	0.6009	Zamanda Yolculuk: Bir Muhasebecinin İnanılmaz Macerası.
	0.5891	Kişisel Yazı.
	0.5876	Zamanlar zor olabilir, ancak iyi seçilmiş teknikler yardımcı olur.
	0.5734	Daha azıyla daha fazlasını elde edin, verimlilik için çabalayın.

4. Sonuç

Bu çalışma, retrieval ensemble yöntemlerinin etkili bir şekilde kullanılarak sorulara doğru cevapların seçilmesinde nasıl bir katkı sağladığı ortaya koymuştur. Deneysel analizler, hem tekil modellerin hem de ensemble yöntemlerin performanslarını detaylı bir şekilde karşılaştırmış ve bu yöntemlerin güçlü yönlerini göstermiştir.

Elde edilen bulgular, ensemble yöntemlerinin tekil modellere kıyasla daha yüksek doğruluk oranlarına ulaşabildiğini göstermektedir. Weighted Cosine Similarity yöntemi, en yüksek top-1 doğruluk oranı ile öne çıkarken, Majority Voting yöntemi top-5 doğruluk oranında üstünlük sağlamıştır. Bu, ensemble yöntemlerinin uygulama amacına bağlı olarak esneklik sunduğunu göstermektedir.

Çalışma, Türkçe dilinde retrieval tabanlı bir sistem üzerinde gerçekleştirilmiş olmasıyla özgün bir katkı sağlamaktadır. Türkçe soru ve cevap çiftleri üzerinde yapılan deneyler, dil bağımsız bir retrieval çerçevesi geliştirilmesine de ışık tutabilir. Daha farklı ensemble yöntemleri kullanılarak farklı veri kümeleri üzerinde yapılacak çalışmalar bu alandaki Türkçe dili özelindeki çalışmaları genişletebilir.

5. Kaynakça

Noyan, M., turkish instructions dataset, https://huggingface.co/datasets/merve/turkish_instructions/discussions, Erişim: 03 Aralık 2024.

Dogan, E., Uzun, M. E., Uz, A., Seyrek, H. E., Zeer, A., Sevi, E., Kesgin, T.T., Yuce M. K., Amasyali, M. F. (2024). Türkçe Dil Modellerinin Performans Karşılaştırması Performance Comparison of Turkish Language Models. arXiv e-prints, arXiv-2404.