



# YAPAY ZEKA (AI) YARIŞMASI

YARIŞMA KURALLARI

## Hedef

Birçok kişiye göre Yapay Zeka (AI), günümüzün en önemli teknolojisidir. Yapay Zeka Yarışması'nın hedefi, bu alandaki ilgiyi artırmak ve katılımcıları hızlı yükselen bu alanda yetkinlik kazanmaya teşvik etmektir. Ayrıca, katılımcıların paylaşımlarıyla Türkçe için eşsiz bir eğitim veri seti oluşturulması hedeflenmektedir. Bu eğitim veri seti, nihai modellerini geliştirmeleri için yarışmacılarla paylaşılacağı gibi, yarışma sonrasında da genel araştırma topluluklarıyla paylaşılacaktır.

## Yarışma Tarifi

Yarışmacılar, Türkçe makaleler temel alınarak Türk İslam Bilim Tarihi alanında test edilmek üzere bir Türkçe Soru Cevaplama sistemi oluşturacaktır.

İngilizce Soru Cevaplama sistemlerinde genel olarak araştırmacılar, cevapları metinlerde mevcut olan soruların kitle-çalışanları tarafından derlenmesiyle oluşturulmuş bir okuma kavrama veri kümesi olan SQuAD<sup>1</sup> kullanırlar.

Türkçede benzer deneme verisi bulmak zordur. Yarışmacılar, bulabildikleri veya oluşturabildikleri her türlü deneme verisini kullanmakta serbesttirler. Bununla birlikte, yarışmacılardan deneme verisi oluşturmanın bir parçası olmaları talep edilecektir. Yarışmacılardan deneme verisi üreterek Yarışma Komitesine teslim etmeleri istenecektir. Tüm yarışmacılardan toplanan deneme verileri, kalitesi teyid edilmesi sonrasında nihai modellerini bu büyük veri kümesi ile deneyebilmeleri için tüm yarışmacılar ile paylaşılacaktır.

Soru Cevaplama testinin bir örneği aşağıdaki şekildedir:

Soru: Türkiye'nin ilk uçak tasarımcısı ve üreticisi kimdir?

Cevap: Vecihi Hürkuş

Soru: Baykar Makina kim tarafından kurulmuştur?

Cevap: Özdemir Bayraktar

---

<sup>1</sup> <https://rajpurkar.github.io/SQuAD-explorer/>

## Yarışma Kuralları

Başvurular TEKNOFEST ISTANBUL Havacılık, Uzay ve Teknoloji Festivali ([www.teknofestistanbul.org](http://www.teknofestistanbul.org)) resmi web sitesinden 15 Mayıs 2018'e kadar gerçekleştirilmelidir.

Takımlar <http://turkiyeteknolojitaakimi.org/icerik/469/-html> linki üzerinden destek talebinde bulunabilirler.

Yarışmacılardan deneme veri kümesine katkı sağlamaları talep edilecektir. Bu amaçla yarışma komitesi, yarışmacı takımlara örnek verilik materyal sağlayacaktır. Her yarışmacı takım, bu materyali veya Türk İslam Bilim Tarihi konusunda başka kaynakları kullanarak deneme veri kümesine en az 500 soru-cevap katkısı sağlamalıdır. Yarışmacılar, yarışmaya devam edebilmeleri için 07.07.2018 tarihine kadar bu talebe karşılık vermelidirler.

500 soru-cevap üzerinde katkı sağlayan takımlar, sağladıkları katkı oranında ek puan alacaklardır. Detaylar için **Puanlama** bölümüne bakınız.

Yarışmacılar, modellerini denemek için istedikleri veriyi kullanabilirler; fakat Yarışma Komitesi 15 Temmuz 2018 tarihinde birleştirilmiş deneme verisini paylaşacaktır. Deneme verisi yeterli görülmediği takdirde, 15 Temmuz 2018 tarihinde takımlara yarışma komitesi tarafından yarışma konusuna benzer konularda ilave verilik materyal sağlanacaktır. Bu materyal üzerinden yarışmacılardan 7 Ağustos 2018 tarihine kadar ilave soru-cevap üretmeleri istenecektir. 15 Ağustos 2018 tarihinde yarışma komitesi, nihai deneme verisini yarışmacılarla paylaşacaktır.

Değerlendirmeye hak kazanabilmek için yarışmacılar, yarışma sonrasında kodlarını açık hale getirmeyi taahhüt ederler. Yarışmacılar açık olması kaydıyla herhangi bir üçüncü parti kodu kullanabilirler.

Nihai kod harici çağrılar kullanmadan, yarışmacının bilgisayarında herhangi bir internet bağlantısı olmaksızın çalışabilmelidir. Değerlendirme gününde yarışmacıların bilgisayarları ile hazır olmaları gerekmektedir.

## Girdi Çıktı Formatları

### a. Veri seti formatı

Veri seti, JSON formatında olup Şekil 1'deki gösterildiği gibi bir düzene sahip olmalıdır. Veri setlerinin JSON düzenine uygunluğu kontrol edilecek olup, uygun olmayan veri setleri kabul edilmeyecektir. JSON nesnelerinin tipleri ve içerdikleri veri bileşenlerinin açıklamaları aşağıdaki gibidir:

- **veri:**
  - tipi: dizi
  - açıklaması: Tüm veri setini içerir
- **başlık:**
  - tipi: metin
  - açıklaması: Metnin konusunu içerir
- **paragraflar:**
  - tipi:dizi
  - açıklaması: Bir metne ait paragrafları içerir
- **paragraf\_metni:**
  - tipi: metin
  - açıklaması: Bir paragrafın içeriğini tutar
- **soru\_cevaplar:**
  - tipi: dizi
  - açıklaması: Bir paragrafa ait soru ve cevapları içerir
- **cevap\_başlangıcı:**
  - tipi: sayı
  - açıklaması: Cevabın paragraf metni içinde sıfırdan başlayarak yukarı sayacak şekilde kaçınıcı kelimedede başladığını gösterir.
- **cevap:**
  - tipi: metin
  - açıklaması: Bir cevabı içerir
- **soru:**
  - tipi: metin
  - açıklaması: Bir soruyu içerir

*b. Model çıktı formatı*

Teslim edilen kod, her biri satır başı karakteri (\n) ile biten soru dosyası okuyabilmeli ve her biri satır başı karakteri (\n) ile biten cevap dosyası üretebilmelidir.

Kodun, satır başı karakteri (\n) ile biten bir soruyu cevaplayabileceği etkileşimli bir kipi de olmalıdır. Etkileşimli kip, büyük harflerle yazılan SONLANDIR komutu ile son bulacaktır.

```
{
  "veri": [{
    "başlık": "Hezârfen Ahmed Çelebi",
    "paragraflar": [{
      "paragraf_metni": "Hezârfen Ahmed Çelebi 17. yüzyılda Osmanlı'da yaşamış olduğu varsayılan Müslüman Türk bilgini. Kendi geliştirdiği takma kanatlarla berberi fizikçi Abbas Kasım İbn Fırnas'tan sonra uçmayı başaran ilk insan olduğu söylene de, kendisi hakkında tek bilgi Seyahatname'de geçmektedir. 1623-1640 yılları arasında saltanat süren Sultan IV. Murad zamanında, uçma tasarısını gerçekleştirdiği ve geniş bilgisinden ötürü halk arasında, Hezarfen olarak anıldığı söylenmektedir. Hezar, Farsça kökenli bir sözcük olup 1000 anlamına gelir. Hezârfen ise \"bin fenli\" (bilimli) yani \"çok şey bilen\" anlamına gelir.",
      "soru_cevaplar": [{
        "cevap_başlangıcı": 34,
        "cevap": "Seyahatname",
        "soru": "Hezârfen Ahmed Çelebi hakkındaki bilgiler nerde geçmektedir?"
      }],
      {
        "cevap_başlangıcı": 64,
        "cevap": "1000",
        "soru": "Hezar ne anlamına gelir?"
      }, {
        "cevap_başlangıcı": 41,
        "cevap": "Sultan IV. Murad",
        "soru": "Hezârfen hangi padişah zamanında uçma tasarısını gerçekleştirmiştir?"
      }
    ]
  }],
  {
    "paragraf_metni": "İlk uçma denemelerinde, Leonardo Da Vinci'nin uçma konusundaki çalışmalarında kendinden çok önce bu konuda deneyler yapan 10. yüzyıl Müslüman Türk alimlerinden olan İsmail Cevheri'den ilham aldığı söylenir. Cevheri'nin bulgularını iyice inceleyen ve öğrenen Çelebi'nin, kuşların uçuşunu inceleyerek tarihi uçuşundan önce hazırladığı kanatlarının dayanıklılık derecesini ölçmek için, Okmeydanı'nda deneyler yaptığı varsayılır.",
    "soru_cevaplar": [{}]
  }
]
},
{
  "başlık": ""
}
]
```

Şekil 1 Örnek veri seti

## Soru-Cevap Hazırlama Yönergesi

Başarılı makine öğrenmesi sistemleri için temiz etiketlenmiş eğitim verisi gerekmektedir. Bu yarışmada kullanılacak veri setinin etiketlenmesi için aşağıdaki metot tavsiye edilmektedir.

- 1) Her grup kendisine etiketlenmek üzere verilen paragraflardan faydalanarak soru ve cevap hazırlayacaktır.
- 2) Her soru için paragraf içinde sorunun cevabı olan kısım **cevap** olarak belirtilecektir.
- 3) Sorular hep aynı kalıp kullanılmadan, metindeki kelimeler çoğunlukla birebir alınmadan, etiketleyicinin kendi kelimeleriyle hazırlanarak çeşitlilik sağlanacaktır. Burada dikkat edilmesi gereken soruların metinde geçen kelimeleri dengeli kullanmasıdır. Aşırı bir şekilde metni takip eden ya da diğer ekstremde cevapla ilgili hiçbir ipucu vermeyen sorular tercih edilmemelidir.
- 4) Kopyala-yapıştır metodu ile üretilen ve çeşitlilik içermeyen sorular pek faydalı olmayacaktır.
- 5) Örnek soru cevaplar:

Konu başlığı: Feza Gürsey

Paragraf: *Feza Gürsey, fizik okumaya lise yıllarında karar vermiştir. Galatasaray Lisesi'ni 1940 yılında birincilikle bitirdikten sonra İstanbul Üniversitesi Fen Fakültesi öğrencisi olmuş, 1944 yılında Fizik-Matematik bölümünden de birincilik ile mezun olmuştur. Millî Eğitim Bakanlığı sınavını kazanarak İngiltere Imperial College'a gitmeye hak kazanmış, burada 1945-1950 yılları arasında Prof. Dr. H. Jones'ın danışmanlığı altında doktora çalışmalarını yapmıştır. Bu dönem içerisinde "Tek boyutlu bir istatistiksel sistem" ve "İki bileşenli dalga denklemleri üzerine" başlıklı iki önemli makale yayımlamıştır. 1951-1957 yılları arasında Cahit Arf'ın desteği ile İstanbul Üniversitesi Tatbiki Matematik Kürsüsü'ne asistan olarak tayin edilmiştir. 1953 yılında "Spinli elektronların klasik ve dalga mekaniği" adlı tezi ile doçent unvanını almış, bir yıl sonra Tatbiki Matematik Kürsüsü'ne doçent olarak atanmıştır.*

*Arzu edilen soru-cevap örnekleri:*

Soru: Feza Gürsey'in doktora çalışmaları sırasında yazdığı makalelerinin başlıkları nedir?

Cevap: "Tek boyutlu bir istatistiksel sistem" ve "İki bileşenli dalga denklemleri üzerine"

Soru: Feza Gürsey doçentliğine hangi kürsüde başlamıştır?

Cevap: Tatbiki Matematik Kürsüsü

Soru: Feza Gürsey doktora çalışmasını hangi bilim adamı gözetiminde tamamlamıştır?

Cevap: Prof. Dr. H. Jones

Soru: Feza Gürsey hangi yıl doçent olmuştur?

Cevap: 1953

Soru: Feza Gürsey hangi liseden mezun olmuştur?

Cevap: Galatasaray Lisesi

Soru: Feza Gürsey hangi yıllar arasında İstanbul Üniversitesi'nde çalışmıştır?  
Cevap: 1951-1957

*İstenmeyen soru örnekleri:*

Metin: **Feza Gürsey**, fizik okumaya lise yıllarında karar vermiştir.  
Soru: **Kim** fizik okumaya lise yıllarında karar vermiştir?

Bu örnekte özne “Feza Gürsey”, soru zamiri olan “Kim” kelimesi ile direk yer değiştirilerek soru üretilmiştir. Soru, sözdizimsel olarak verilen metni yakından takip ediyor.

Metin: **1951-1957 yılları arasında** Cahit Arf'ın desteği ile İstanbul Üniversitesi Tatbiki Matematik Kürsüsü'ne asistan olarak tayin edilmiştir.  
Soru: **Hangi yıllar arasında** Cahit Arf'ın desteği ile İstanbul Üniversitesi Tatbiki Matematik Kürsüsü'ne asistan olarak tayin edilmiştir?

Metin: **1944 yılında** Fizik-Matematik bölümünden de birincilik ile mezun olmuştur.  
Soru: **Hangi yılında** Fizik-Matematik bölümünden de birincilik ile mezun olmuştur?

Bu örneklerde de yine “Hangi” soru sıfatı, metindeki zaman zarflarıyla yer değiştirilmek suretiyle verilen metne sözdizimsel olarak çok benzer sorular üretilmiştir.

Metin: Feza Gürsey, fizik okumaya **lise yıllarında** karar vermiştir.  
Soru: Feza Gürsey, fizik okumaya **hangi zaman diliminde** karar vermiştir?  
Cevap: lisede

Bu örnekte cevap, metin içinde bir kelime aralığını içermiyor. Cevap metinden türetilmiş. Doğru cevabın “lise yıllarında” olması gerekiyor.

Metin: Devrin en zengin iş adamı olan Demirağ., 1936 yılında uçak fabrikasına kurma girişimine başladı. O yıllarda ordunun uçak ihtiyacı **halktan ve zengin isadamlarından toplanan bağışlarla** karşılanmaktaydı.  
Soru: Demirağ uçak fabrikasını kurma girişimine başladığı yıllarda ordunun uçak ihtiyacı **nasıl** karşılanıyordu?  
Cevap: bağışlarla

Bu örnekte cevap muğlak kalıyor. Doğru cevabın metinde altı çizili kelime aralığı olması gerekiyor.

## **Puanlama**

Puanlama, model performansı ve veri toplama olarak iki kısma ayrılacaktır. Model performansında soruya verilen cevabın tam doğruluğu kriteri aranacaktır. Yaklaşık cevaptan kısmi puan alma söz konusu değildir.

Model performansı, 80 puan üzerinden değerlendirilecek olup, en yüksek tam doğruluk oranını elde eden takım (eşitlik durumunda en yüksek tam doğruluk oranını elde eden her bir takım) 80 tam puan alacaktır. En yüksek tam doğruluk oranının altında bir doğruluk oranı elde eden takımların puanı ise şu şekilde hesaplanacaktır. Takımın elde ettiği tam doğruluk oranı, elde edilen en yüksek tam doğruluk oranına bölünecektir. Çıkan sonuç 80 tam puan ile çarpılarak takımın model performansından elde ettiği puan belirlenecektir. Örneğin, yarışan modeller içinde elde edilen en yüksek tam doğruluk oranı %60 olsun. Bu modelin sahibi takım, 80 tam puan kazanacaktır. Başka bir takımın modeli %55 tam doğruluk oranı elde ediyor olsun. Bu takımın kazanacağı model puanı ise  $(55 / 60) \times 80 = 73.33$  olacaktır.

Veri toplama puanı ise 20 puan üzerinden değerlendirilecek olup, en fazla veri katkısı yapan takım (eşitlik durumunda en fazla veriyi sağlayan her bir takım) 20 tam puan alacaktır. En fazla veri sayısının altında katkı sağlayan takımların puanı ise şu şekilde hesaplanacaktır. Takımın sağladığı veri sayısı en fazla sağlanan veri sayısına bölünecektir. Çıkan sonuç 20 tam puan ile çarpılarak takımın veri toplama puanı belirlenecektir. Örneğin en fazla veri sağlayan takımın sağladığı veri sayısı 1000 olsun. Bu takım 20 tam puan kazanacaktır. Diğer bir takım ise 500 veri katkısı yapmış olsun. Bu takımın veri puanı ise  $(500 / 1000) \times 20 = 10$  olacaktır.

Bir takımın toplam puanı model performans puanının veri toplama puanıyla toplanması sonucu elde edilecektir.

Puan eşitliği durumunda model puanı daha çok olan takım öne geçecektir.

Yarışmacılar, Eylül ayındaki (gün daha sonra duyurulacaktır) değerlendirme gününde USB'den girdi olarak verilecek dosya ile programlarını çalıştıracaklardır.

Nihai değerlendirmeden önce sistemlerini doğrulayabilmeleri için yarışmacılara 15 Temmuz 2018 tarihinde örnek soru kümesi sağlanacaktır.

## **Minimum Başarı Kriteri**

Benzer İngilizce yarışmalarda 80% üzeri başarı sağlanabilmekte, ve sonuçlar insan performansına hızla yaklaşmaktadır. Yarışmacılarımız, kendi üretecekleri, daha az olgunlaşmış bir veri kümesi ile çalışacaklarından sonuçlar aynı



seviyelere ıkamayabilir. Öte yandan, bu yarışma belirli bir konu etrafında yapıldığından, İngilizce yarışmalardan daha kolay olacaktır. En az %55 tam doğruluk oranı bu yarışma için minimum başarı kriteri olacaktır. Bu şartı sağlamayan yarışmacılar ödüle hak kazanamazlar.

## **Ödüller**

Puanlamada ilk üç sıraya giren ve en az %55 tam doğruluk uyumu yakalayan takımlar kaynak kodlarını github.com'a net kurulum talimatları ile yüklemelidir. Farklı bilgisayarlarda sonuçların tekrarlanabileceğinin teyit edilmesi sonrasında takım ödüle layık görülür. Aksi durumda takım elenir ve bir sonraki en yüksek puanlı yarışmacıya geçilir.

Yarışma sonrasında, işbu şartnamede belirtilen başarı kriterini başarıyla yerine getirerek ödül sıralamasına giren yarışmacılar derecelerine göre aşağıdaki gibi ödüllendirilecektir. Ödüller takımlara verilecek olup, bireysel ödüllendirme yapılmayacaktır.

Birincilik Ödülü: 30.000 TL

İkincilik Ödülü: 20.000 TL

Üçüncülük Ödülü: 10.000 TL

## GENEL KURALLAR & DÜZENLEMELER

- Her takımın yetkili kişilerinin ilgili hakeme itiraz hakkı vardır. İtirazlar sonradan yazılı olarak verilmek kaydıyla sözlü olarak da yapılabilir. Sözlü olarak yapılan itirazlar en geç 24 saat içerisinde yazılı hale getirilir. Her halükarda yazılı olmayan itirazlar dikkate alınmayacaktır. Yapılan itirazlar hakem heyeti tarafından incelenerek 24 saat içerisinde karara bağlanır.
- Her bir yarışmacı yarışırken gerekli emniyet tedbirlerini almak ve çevresine karşı kendisinden beklenen özeni göstermekle yükümlüdür.
- Yarışmacıların 3. kişilere verdiği zararlardan T3 Vakfı ve organizasyon yetkilileri sorumlu değildir.
- T3 Vakfı ve organizasyon komitesi, adil sonuçlar doğurabilmesi açısından yarışmaların objektif kriterler içerisinde gerçekleşmesi, yarışmacıların her türlü ihtiyaçlarının daha iyi karşılanabilmesi, emniyet tedbirlerinin sağlanması ve yarışma şartlarının işlerlik kazanabilmesi için işbu şartnamede her türlü değişiklik yapma hakkını saklı tutar.
- T3 Vakfı ve organizasyon komitesi, yarışmalara başvuru sürecinin ardından gerçekleştirilecek değerlendirmeler sonucunda, yarışmalara katılmak için gerekli teknik bilgi ve becerilere sahip yeterli başvuru olmaması durumunda yarışmaları iptal etme hakkını saklı tutar.
- TEKNOFEST İSTANBUL Havacılık, Uzay ve Teknoloji Festivali Güvenlik ve Emniyet Şartnamesi tüm yarışmacılara, heyetlerine ve ilgili kişilere tebliğ edilir. Organizasyon kapsamında yarışacak bütün takımlar, TEKNOFEST İSTANBUL Havacılık, Uzay ve Teknoloji Festivali Güvenlik ve Emniyet Şartnamesi'nde yarıştıkları yarışma özelinde belirtilen güvenlik şartlarını sağlamakla yükümlüdür. Bu bakımdan, söz konusu emniyet talimatında yer alanlar haricinde, kullanılan sistemlerden kaynaklı ilave tedbirlerin alınması yarışmacıların sorumluluğundadır.
- T3 Vakfı ve organizasyon komitesi, TEKNOFEST İSTANBUL Havacılık, Uzay ve Teknoloji Festivali Güvenlik ve Emniyet Şartnamesi'nde belirtilen koşulları sağlamadığını tespit edilen takımları, organizasyonun güvenli bir ortamda gerçekleşebilmesi adına yarışma dışı bırakma hakkını saklı tutar. Yarışmacıların, heyetlerinin ve ilgili kişilerinin yarışmalar esnasında doğan ihlalleri sonucunda oluşabilecek zararlardan T3 Vakfı ve organizasyon yetkilileri sorumlu değildir.
- Yarışmaya katılma hakkı kazanan tüm takımlara sertifika verilecektir.

**Türkiye Teknoloji Takımı Vakfı işbu şartnamede her türlü değişiklik yapma hakkını saklı tutar.**