

TÜRKÇE DOĞAL DİL İŞLEME ONLINE

YARIŞMA ŞARTNAMESİ

Yarışmanın Adı: Türkçe Doğal Dil İşleme Online Yarışma Etkinliği

Koordinatör: Tuba Öztepe - Bilişim Vadisi Proje Yönetim Ofisi Müdürü

Yarışmanın Çalışma Grubu

- Kemal Can Kara - Kariyer.net
- Murat Tan - TÜBİTAK BİLGEM YTE
- Ramazan Gökay - TÜBİTAK BİLGEM YTE
- Enes Bulut - Garanti BBVA Teknoloji
- Alara Dirik - University of Glasgow
- Yusuf Yiğit - TRT
- Necmettin Çarkacı – TRT World
- Emir Karşıyakalı - Kommunity.com
- Emrah Budur - Garanti BBVA Teknoloji

Yarışmanın Amacı

Türkiye Açık Kaynak Platformu tarafından Türkçe Doğal Dil İşleme konusunda farkındalık oluşturmak amacıyla online yarışma düzenlenmektedir. İlgili yarışma süresince katılımcılardan;

- Veri Üretme
- Açık Kaynak Araçları Kullanma
- Veri Setleri Geliştirme
- İşaretsiz Verileri İşaretli Hale Getirme gibi konularda faaliyet göstermeleri istenecektir.

Yarışmanın Hedef Kitlesi

Yarışma kapsamında belirlenen konularda ya da verilen vakalar üzerinden proje üretmek isteyen 18 yaş üstü herkes ilgili yarışmaya katılım sağlayabilir.

Yarışmanın Konusu

Online yarışmanın ana konusu Türkçe Doğal Dil İşleme'dir. İlgili konuda aşağıda belirlenen alt başlıklarda projeler oluşturulacaktır:

- Metin düzeltme (normalization)
- Tanımlanan veriler üzerinden soru cevap işlemleri
- Covid 19 a özel (tahmin/veri anlamlandırma/bilgi çekme) çalışmalar
- Doğal dil anlama(natural language understanding)

- Doğal dil üretimi (natural language generation)
- Türkçe Doğal Dil İşleme’de Bağlamsal Anlama Teknolojisi (BERT vs.)
- Optik karakter tanımlama (OCR) - Türkçe Destek
- İlan metinlerinde kişinin / kurumların isterlerini bulma
- Açık kaynak hedef araçların Türkçe içinde aynı performans seviyesine getirilmesi
- Mevcut kulvarlar üzerinden, veri setleri paylaşılarak, bunlar üzerinden çözümlerin geliştirilmesi
- Temel problemler üzerinden ilerleme
- Türkçe Karakter Tanımlama
- NLP Kütüphanelerinin geliştirilmesi, veri seti oluşturma, (veri etiketleme araçlarında eksikler mevcut)

Yarışmanın Kapsamı

1. Etkinlikten en az 1 ay önce duyurular başlatılacaktır.
2. Düzenlenecek olan yarışma online ortamda gerçekleştirilecektir.
3. Etkinlik süresince ihtiyaç duyulacak altyapı desteği Kommunity tarafından sağlanacaktır.
4. Etkinlik süresince yarışmacılara destek olması planlanan mentörlerin ve jüri üyeleri önceden belirlenecektir.
5. Yarışmacılar ile paylaşılacak veriler standartlaştırılacaktır.
6. Datayı paylaştıktan sonra yarışmacıların projeyi nasıl geliştirecekleri hususunda serbest bırakılacaktır.
7. Problem genel kitleye hitap etmeli, ortaya çıkan sonuç genel kitle tarafından kullanılmalı.
8. Yarışma süresince katılımcıların belirlenen alt başlıklar hakkında ya da konu ile ilintili farklı alanlarda proje geliştirmesi istenmektedir.

Yapılması Planlanan İşler

- Yarışmanın başvuru süresinde, tüm katılımcılardan proje fikirlerini anlattıkları basit bir sunum hazırlamaları talep edilmektedir. İletilen sunumlar üzerinden ön eleme yapılacaktır.
- Jüri ve Mentör Listesi;
 - Kemal Can Kara - Kariyer.net
 - Murat Tan - TÜBİTAK BİLGEM YTE
 - Ramazan Gökay - TÜBİTAK BİLGEM YTE
 - Enes Bulut - Garanti BBVA
 - Alara Dirik - University of Glasgow
 - Emrah Budur - Garanti BBVA
 - Yusuf Yiğit - TRT
 - Necmettin Çarkacı – TRT World
 - Web sitesi üzerinde yarışma kapsamında mentörlük yapmak isteyen kişiler için “Mentörlük Başvuru Formu” eklenecektir.

- Katılımcılar ile mentörler zoom, teams gibi platformlar üzerinden görüşme yapacaktır. Google form üzerinden ortak dosya oluşturulup mentörlerin uygun zamanlarının belirtildiği bir takvim oluşturulacaktır.
- Mentörler aynı zamanda jüri üyesi olacaktır. Buna ek olarak jüri üyesi olarak farklı isimler eklenebilir.
- Katılımcıların başvuruları acikhack.com web adresi üzerinden alınacaktır. Başvuru kısmına gelen yarışmacı başvuru yapmak için Kommunity.com'a üye olacak ve buradan başvurusunu gerçekleştirecektir. Başvuru esnasında sunum talep edilmektedir. (www.acikhack.com)
- Projelerin Değerlendirme Süreçleri
 - Yarışmacılar için projelerini sunacakları bir sunum formatı belirlenecektir.
 - Github üzerinden teknik bir blog hazırlamaları talep edilecektir.
 - GitHub yol gösterici doküman hazırlanıp, web sitesinde paylaşılacaktır.
- **Kazanan Ekibe Ödüller**
 - **Birinci olan ekip 1 yıl boyunca Bilişim Vadisi Kuluçka Merkezi'nde ücretsiz olarak ofis imkanı elde edecektir. Ayrıca ekibin tamamına Macbook Pro dizüstü bilgisayar hediye edilecektir.**
 - **İkinci ve üçünü olan ekibe Macbook Air dizüstü bilgisayar hediye edilecektir.**
- Acik Seminer'de verilen, Türkçe Doğal Dil İşleme ve Python konuları ile ilgili seminerler acikhack web sitesine link olarak bırakılacaktır. Bu sayede yarışmacılarda farkındalık artırılabilir.

Proje Değerlendirme kriterleri *****

- Türkçe Doğal Dil İşlemeye Katkısı (40)
 - Kodlama açısından yeterlilik
 - Topluluğa sağlanan katkı
 - Projenin yenilik seviyesi
 - State of Art
 - Mevcut çözüm üzerinden ilerleme durumunda teknik çözüm ve fikrin değeri
 - Proje kapsamında sağlanan katkı
 - Kullanım olarak sunulan katkı
 - Kullanılabilecek bir çıktı oluşturulması
- Yaratıcılık - İnovasyon (40)
 - Sosyal Etki
 - İş Değeri yaratmak
- Sunum performansı (10)
- Dokümantasyon (10)***** - gereksinimler mutlaka yazılmalı
- Somut cevaplar verilebilir bir puanlama sistemi oluşturulmuştur.
 - Yarışma sonucunda katılımcılara verilen puanlar açıklanmayacaktır. İtiraz durumunda açıklanabilir.

Yarıřmaya Bařvuru Kriterleri

- Yarıřmacılar bařvuru formunu doldururken girdikleri tüm bilgilerin doęru olduęunu kabul eder.
- Yarıřmaya 18 yař üstü herkes bařvuruda bulunabilir.
- Katılımcılar, bařvuru sürecinde; kendilerini, takım arkadaşlarını ve proje fikirlerini anlatan bir sunum hazırlamak zorundadır. Söz konusu sunumu Google Drive ya da Dropbox gibi bir bulut alanına yükleyip, söz konusu alanın linkini bařvuru formunun en alt kısmında yer alan alana girmek zorundadır.
- Yarıřmaya katılım bireysel ya da takım halinde olabilir. Takımlar en fazla dört katılımcıdan oluşabilir.
- Yarıřma kapsamında hazırlanacak sunumda takım üyelerinin tamamının proje kapsamında görev tanımları yapılmalıdır.
- Yarıřma süresince geliştirilecek kodların GitHub'da paylaşımı zorunludur. Bu bağlamda tüm yarıřmacılar projeleri ile ilgili gerekli dokümantasyonu yapmak zorundadır.
- Yarıřmacılar, Türkçe Doğal Dil İşleme ile ilgili yaptıkları çalışmalarını Platform ile paylaşmalıdır.
- Yarıřmacılar proje geliştirme süresince GitHub üzerinden para ile satın aldıkları kütüphaneleri kullanılamaz.
- Yarıřmacılar, yarıřmaya mevcut durumda üzerinde çalıştıkları projelerini sunamaz.
- Yarıřma süresince, yarıřmacıların daha önce başka yerde yayınlanmış ya da satılmış projeleri kullanılamaz.

Yarıřmanın Çıktıları	<ul style="list-style-type: none">• TRT, Garanti BBVA, TÜBİTAK gibi kurumların mentörlük destekleri• Yarıřmacılar tarafından geliştirilecek projeler• Türkçe Doğal Dil İşleme ile ilgili ülkemizde farkındalık yaratılması
Zaman Planı:	<ul style="list-style-type: none">• Duyuru Tarihi - 8 Haziran 2020• Kayıt alma süreci - 21 Haziran 2020• Sunumların Deęerlendirilmesi Süreci – 22-26 Haziran 2020• Bilgilendirme Toplantısı - 20 Haziran 2020• Online yarıřma bařlangıç - 01 Temmuz 2020<ul style="list-style-type: none">○ Haftalık mentör görüşmeleri• Online yarıřma bitiş - 30 Temmuz 2020• Deęerlendirme süreci ve dereceye girenlerin belirlenmesi – 01-30 Ağustos 2020• Dereceye girenlerin ödülleri TEKNOFEST'te verilmesi (Eylül 2020)