



LISAN-I TRKİ

Trke Doęal Dil İřleme Yarışması



Lisan-ı Türki Nedir?

Takımımız 2020 yılının sonlarında Türkçe doğal dil işleme kaynaklarını artırma gayesiyle kurulmuştur. Doğal dil işleme, makine öğrenmesi alanında en çok ele alınan konu olmasının yanı sıra, çalışmalar büyük oranda İngilizce dili üzerine gerçekleştirilmektedir. Yapısal olarak İngilizce dili ile Türkçe dili çok farklı olduğu için o dilde yapılmış bir çalışmayı direkt olarak Türkçe diline uygulama şansımız olmamaktadır. Bu da bizi Türkçe dili için yeni teknik ve yöntemler aramaya itmektedir. Takımımız da bu motivasyon ile doğal dil işleme alanının en popüler başlıklarını (metin özetleme, isim-varlık çıkarımı, duygu analizi) Türkçe dili için uyarlamaya ve bu alandaki kaynak sayısını arttırmayı hedeflemektedir.



Üye Bilgisi

- **Oğuzhan Şahin**
 - 2020-2021 yılı arasında Muğla Sıtkı Koçman Üniversitesi, Bilgisayar Mühendisliği 4.sınıf öğrencisi olarak öğrenim görmektedir.
 - Bu takımın ilk ve tek üyesidir. Bu sebeple bütün iş yükünü üstlenmektedir.
 - Yaklaşık 1.5 yıldır makine öğrenmesi, yaklaşık 1 yıldır ise doğal dil işleme ile ilgilenmektedir.

Proje Fikri: Türkan



LISAN-I TÜRKI



Motivasyon

Günümüzde artık gündemdeki haberler gazetelerden ziyade sosyal medya üzerinden takip edilmektedir. Özellikle Twitter üzerinden haber siteleri güncel olarak kendi profilinde bu haberleri paylaşmaktadır. Ancak bu haber sitelerinin çoğunda clickbait (tıklama tuzağı) yapılmaktadır. Site içerisinde geçirilen vakti artırmak için ulaşılmak istenen bilgi haber siteleri tarafından doğrudan verilmez. Yani kelime yığınları arasında istenen bilgiye ulaşmak epey bir vakit almaktadır. Bu da beraberinde okuma isteğini azaltıp zaman kaybına sebebiyet vermektedir. Bunun yanı sıra, bütün haberleri okuyacak vakit her zaman olmayabilir, bu durumda verilmek istenen mesajı doğrudan görmek zaman kaybını önlemektedir. Buna çözüm olarak Türkan'ı öneriyoruz. Türkan, önceden eğitilmiş Türkçe BERT question-answering (soru-cevap) modelini kullanan bir Twitter botudur.

Yöntem ve Teknikler



Resim 1. Örnek Tweet

Resim 1. deki örnek paylaşımda görüldüğü üzere bir soru cümlesi ardından o soru cümlesinin cevabını içeren haber linki mevcuttur. Bu sorunun cevabına ulaşmak için ilgili habere gidip gereksiz detayları okumak yerine Türkan, ilgili tweetteki soruyu ve haber linkini analiz eder, cevabı bu tweeti alıntılıyarak paylaşır.

Yöntem ve Teknikler



Resim 2. Örnek Tweet

Ancak buradaki problem, Resim.2’de de görüldüğü üzere bir tweet içerisinde her zaman soru cümlesi olmayabilir. Buna çözüm olarak, ilgili cümlenin morfolojik yapısı analiz edilerek soru cümlesine dönüştürülecektir. “Kulüpler Birliği’nin yeni başkanı belli oldu.” cümlesi “Kulüpler Birliği’nin yeni başkanı kim oldu?” sorusuna dönüştürülecektir.