

Türkiye Açık Kaynak Platformu
Online Yarışma Programı

Türkçe Doğal Dil
İşleme

Türkiye



Turkey Open

Açık KaynakSource Platform

Platformu«/»www.turkeyopensourceplatform.com

ACIKHACK

Açık Kaynak Hackathon Programı

www.acikhack.com

- THE TAGGERS -

Universal Analyzer

www.turkiyeacikkaynakplatformu.com



Türkiye



Turkey Open

Açık KaynakSource Platform

Platformu«/»www.turkeyopensourceplatform.com

ACIKHACK

Açık Kaynak Hackathon Programı

-Merve Din
-Ecem Bölük
-Dilara Kanalıcı

www.turkiyeacikkaynakplatformu.com





Morfolojik olarak çok zengin olduğu için Türkçe için pos etiketleyici yoktur. Bu nedenle, bir pos etiketleyici oluşturmak yerine bir morfolojik analizör geliştirdik. Evrensel bir şey bulmak, diğer dillerle ortak paydada buluşmak istedik. Projemiz hakkında ne olacağına karar verdiğimizde, bunu yapmanın modern bir yolunu aramaya başladık. Yaklaşımımız çok fazla insan çabası gerektirmez. word ve character level embedding yerleştirmeyi kullanıyoruz. Word embedding, bir kelimenin



Türkiye



Turkey Open

Açık KaynakSource Platform

Platformu«/»www.turkeyopensourceplatform.com

ACIKHACK

Açık Kaynak Hackathon Programı

semantic araçlarını öğrenebilir ve character embedding, synthetic özellikleri yakalayabilir. Bunlardan yararlanarak universal bir tag düzeneği ve pos-tagging işlemi yapıldı.

www.turkiyeacikkaynakplatformu.com



Türkiye



Turkey Open

Açık KaynakSource Platform
Platformu«/»www.turkeyopensourceplatform.com

ACIKHACK

Açık Kaynak Hackathon Programı

www.turkiyeacikkaynakplatformu.com





- İŞ AKIŞI -

The Different POS Tagging Techniques

There are different techniques for POS Tagging:

1. **Lexical Based Methods** — Assigns the POS tag the most frequently occurring with a word in the training corpus.
2. **Rule-Based Methods** — Assigns POS tags based on rules. For example, we can have a rule that says, words ending with “ed” or “ing” must be assigned to a verb. Rule-Based Techniques can be used along with Lexical Based approaches to allow POS Tagging of words that are not present in the training corpus but are there in the testing data.
3. **Probabilistic Methods** — This method assigns the POS tags based on the probability of a particular tag sequence occurring. Conditional Random Fields (CRFs) and Hidden Markov Models (HMMs) are probabilistic approaches to assign a POS Tag.
4. **Deep Learning Methods** — Recurrent Neural Networks can also be used for POS tagging.



Türkiye



Turkey Open

Açık KaynakSource Platform

Platformu«/»www.turkeyopensourceplatform.com

ACIKHACK

Açık Kaynak Hackathon Programı

– Ecem Bölük:

Veriseti oluşturma

– Dilara Kanalıcı:

Veriseti oluşturma

www.turkiyeacikkaynakplatformu.com



Türkiye



Turkey Open

Açık KaynakSource Platform
Platformu«/»www.turkeyopensourceplatform.com

ACIKHACK

Açık Kaynak Hackathon Programı

— Merve Din: Model geliştirme çalışması

www.turkiyeacikkaynakplatformu.com

