

Doğal Dil İşleme'de (NLP) **Zemberek**, Türkçe metinler üzerinde çeşitli dil işleme görevlerini gerçekleştirmek için geliştirilmiş açık kaynaklı bir yazılımdır. Bu proje, özellikle Türkçe'nin yapısal özelliklerini dikkate alarak metin işleme, analiz ve dil bilgisi kontrollerini otomatikleştirme amacıyla oluşturulmuştur. Zemberek projesi, Türkçe'nin yanı sıra benzer yapıya sahip diğer diller için de uyarlanabilir.

Zemberek'in Başlıca Özellikleri:

1. **Morfolojik Analiz:** Türkçe kelimelerin ek yapısı oldukça karmaşıktır. Zemberek, kelimeleri kök ve eklerine ayırarak analiz yapar. Örneğin, "gittim" kelimesini "git" (kök) + "tim" (şahıs eki) şeklinde bölebilir.
2. **Heceleme ve Kök Bulma:** Bir kelimenin kökünü ve eklerini ayırma işlevinin yanında hecelemesini de doğru bir şekilde yapabilir.
3. **Yazım ve Dilbilgisi Denetimi:** Yanlış yazılmış veya hatalı dil bilgisi kurallarına göre oluşturulmuş cümleleri tespit edip öneriler sunabilir.
4. **Kelime Üretimi:** Verilen bir kökten anlamlı kelimeler türeterek dildeki kullanım potansiyelini analiz edebilir.
5. **Doğal Dil Anlama:** Özellikle anlam çıkarımı ve metin sınıflandırma gibi NLP görevlerinde Türkçe için temel işleme birimleri sağlar.

Zemberek Kullanım Örnekleri:

- **Arama Motorları:** Türkçe arama sorgularını daha doğru anlamlandırmak ve kullanıcının arama amacına uygun sonuçlar döndürmek için.
- **Sosyal Medya Analizi:** Türkçe metinlerin duygu analizi, konularına göre sınıflandırılması veya trend analizleri için.

- **Chatbot ve Asistanlar:** Türkçe dil desteği sağlamak ve doğru yanıtlar üretmek amacıyla.

Zemberek, Türkçe'nin yapısal özelliklerine göre özelleşmiş güçlü bir araç olup, Python veya Java gibi çeşitli programlama dilleriyle entegre edilerek kullanılabilir.

Örnek çalışma:

```
import zemberek.morphology.TurkishMorphology;
import zemberek.morphology.analysis.SingleAnalysis;
import zemberek.morphology.analysis.WordAnalysis;

public class ZemberekExample {
    public static void main(String[] args) throws Exception {
        // Zemberek'in Türkçe morfolojik analizini başlatıyoruz
        TurkishMorphology morphology = TurkishMorphology.createWithDefaults();

        // Analiz edilecek kelime
        String input = "kitaplığı";

        // Morfolojik analiz
        WordAnalysis results = morphology.analyze(input);

        // Sonuçları yazdırma
        for (SingleAnalysis result : results) {
            System.out.println("Kök: " + result.getStem() + ", Ekler: " + result.getMorphemes());
        }
    }
}
```

Kök: kitap, Ekler: [Noun, A3sg, P3sg, Nom]

GÜNLÜK HAYATTAN ÖRNEKLER:

Zemberek'in günlük hayattan örnekleri, özellikle Türkçe dil işleme (NLP) görevleriyle ilgili alanlarda karşımıza çıkar. İşte bazı pratik örnekler:

- 1. Otomatik Düzeltme ve Dil Bilgisi Kontrolü:** Cep telefonlarındaki otomatik düzeltme ve dil bilgisi kontrol özellikleri Zemberek gibi dil işleme araçlarıyla yapılabilir. Örneğin, bir mesaj yazarken yapılan yazım hatalarını otomatik olarak düzeltebilir veya öneriler sunabilir.
- 2. Sosyal Medya Analizi:** Zemberek, sosyal medya gönderilerini analiz ederek duygu (pozitif, negatif, nötr) tespiti yapmada kullanılır. Bir şirketin sosyal medya hesabında kullanıcı yorumlarının analizini yapmak, müşteri memnuniyetini anlamak için oldukça kullanışlıdır.
- 3. Arama Motoru İyileştirmesi:** Arama motorlarında yapılan sorguların köklerine inmek, arama motorlarının daha doğru sonuçlar sunmasına yardımcı olur. Örneğin, kullanıcı "kitaplar" diye arama yaptığında, Zemberek kök analizini kullanarak "kitap" terimine odaklanabilir, böylece tekil ve çoğul biçimlerini aynı anda ele alır.

4. Metin Sınıflandırma ve Filtreleme: Çeşitli platformlarda veya e-posta hizmetlerinde, Zemberek küfür ya da spam tespiti için metinleri analiz edebilir. Örneğin, uygunsuz içerikleri tespit etmek için kelime kökü ve ek analizleri yapılır.