

# YILDIZ TEKNİK ÜNİVERSİTESİ BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ DOĞAL DİL İŞLEME DERSİ ÖDEV 1: REGEX FATMA ZEHRA ÇETİN 17011068

### 1. Yazdığım RegEx

([A-zÇĞİÖŞÜÂçğiöşü0-9"\.\-]\*MAHALLES?[İI]?|[A-zÇĞİÖŞÜÂçğiöşü0-9\.\-"]\* MA?H?\.?)?([\."A-zÇĞİÖŞÜÂçğiöşü0-9-,]\*CADDES?[İI]?|[\."A-zÇĞİÖŞÜÂçğiöşü0-9-,]\*CA?D\.?)?([\."A-zÇĞİÖŞÜÂçğiöşü0-9\/]\*SOKAK?Ğ?I?|[\."A-zÇĞİÖŞÜÂçğiöşü0-9\/]\*SOK\.?|[\."A-zÇĞİÖŞÜÂçğiöşü0-9\/]\*SK\.)?(\*NO.\*)?(.\*)?([A-zÇĞİÖŞÜÂçğiöşü]\*)(\/)([A-zÇĞİÖŞÜÂçğiöşü]\*)

Kelimeleri mahalle, cadde, sokak, no, il ve ilçe olarak ayrı ayrı gruplamayı amaçladım.

# 2. Örnek Çıktı

```
ATATÜRK BULVARI 32 C FATİH/İSTANBUL

KÜLTÜR SOK. NO:2 BAHÇELİEVLER/İSTANBUL

MERKEZ MAH. ALTIN YILDIZ SİTESİ SİTE SOK. NO:8 BAHÇELİEVLER/İSTANBUL

SOFULAR MAH. PARK İÇİ MACAR KARDEŞLER CAD. NO:1/1 FATİH/İSTANBUL

GAZİ MAH. İSMET PAŞA CAD. NO:109/A GAZİOSMANPAŞA/İSTANBUL

MEHMET AKİF MAH. TEPE ÜSTÜ GÜR DEMİREL C29 KÜÇÜKÇEKMECE/İSTANBUL

HÜRRİYET MAH. KUYU K.K NO:2 E BAĞCILAR/İSTANBUL

ERENKÖY MAH. ŞEMSETTİN GÜNALTAY CAD. ÖĞRETMEN HAYRULLAH DURAĞI KADIKÖY/İSTANBUL

KAYIŞDAĞI MAH. USLU CAD. NO:91/B ATAŞEHİR/İSTANBUL

MAKASÇILAR CAD. NO:59/10 FATİH/İSTANBUL

CAKMAK MAH. OKUL SOK. NO:67/B ÜMRANİYE/İSTANBUL

İSTASYON MAH. YENİ DOĞAN MEVKİİ ŞEN DURAK İST. CAD. NO:17 KÜÇÜKÇEKMECE/İSTANBUL
```

Her adres satırında mahalle, cadde, sokak, numara, ilçe ve il gibi kısımlar farklı gruplara dahil olmuştur.

### 3. Başarı Oranı

6829/6831 = %99'dur. 6829 adresi bulmuştur ve doğru ayırmıştır. Bulduğu adresleri eksiksiz bir biçimde bulmuştur. Tüm adres satırları işaretlense de bazı uç durumlarda mahalle, sokak, cadde ayrımlarında hata olabilir.

# 4. Eksik/Yanlış Bulduğu Adresler

2 adresi doğru ayıramadığını gözlemledim.

Gözüme çarpan bir hatalı veri şu olmuştur:

```
ŞİRİNEVLER MAH MAHMUTBEY CAD NO:80/A BAHÇELİEVLER/ İSTANBUL
```

Mahalle ayrılırken MAH kelimesi üzerinden eşleme yaptığım için Mahmutbey'in ilk hecesi mahalle kısaltması olarak işaretlenmiştir.

```
DEMİR KAPI MAH. VELİ OĞLU CAD.NO:127 BAĞCILAR/ İSTANBUL
```

Bu adresin mahalle ve caddesi doğru ayrılmamıştır. Sebebini çözemedim.

## 5. Bulamadığı Adresler

Yoktur.