



Doğal Dil İşlemeye Kavramsal Bir Bakış Ödevi

Dersin Yürütücüsü: *Banu Diri*

Öğrencinin;

Adı: Muhammed Kayra

Soyadı: Bulut

Numarası: 23501901

Öncelikle bütün adreslere bakıp genel niteliklerini incelediğimde, çoğunun sonunda /İstanbul olduğunu fark ettim.

Bu ifadeleri `(.*[/]\s*[İi](stanbul|STANBUL))` regex kalıbı ile bir araya getirdim. Ayrıca `(T.*NO:.*)` i ifadesini de başka bir durum olarak ele aldım.

Başlangıç ifadesi aşağıdaki gibidir. `(T.*NO:.*)(.*[/]\s*[İi](stanbul|STANBUL))`

Sonuçlara bakarken, ilk çözümümdeki regex'in yetersiz olduğunu fark ettim. 'adresteki işler tamamlandıktan sonra Yıldırım Mah. Kemal Paşa Sok. No:8 BAYRAMPAŞA/İSTANBUL' cümlesinde, ifadenin tamamı adres olarak algılanıyordu. Ayrıca, adreslerin 'YENİBOSNA' gibi başlangıç kısımları da eşleşmiyordu. Bu nedenle, özelleştirme yapmaya karar verdim. Bu ve benzeri cümleleri incelediğimde, adreslerin büyük harfle başlaması gerektiğine karar verdim. İfadenin başındaki `*` regex'in sorunlu olduğunu gördüğüm için değiştirmeye başladım. Büyük harfle başlayan kelimeleri (tek harfliler dahil) kapsayan regex'i `[A-ZĞÜİÖÇŞ 0-9]` olarak yazdım. Bunların tekrar edebilmesi için, sonunda boşluk veya tab karakterleri de olabileceğini düşünerek regex'i genişlettim. Sorgu bu haliyle aşağıdaki haline ulaştı.

`([A-ZĞÜİÖÇŞ 0-9][\t]*)+`

Daha sonra, adreslerin içinde `.`, `:`, `-`, `/` gibi karakterlerin de bulunabileceğini göz önünde bulundurarak, regex'in ikinci kısmı olan **boşluk** bölümünü `[.-:/]` şeklinde genişlettim. Sorgu bu haliyle aşağıdaki haline ulaştı.

`(T.*NO:.*)([A-ZĞÜİÖÇŞ0-9][\.:\/]*[\t]*)+([/]\s*[İi](stanbul|STANBUL))`

Örnekleri incelediğimde, 'Yıldırım Mah. Kemal Paşa Sok. No:8 BAYRAMPAŞA/İSTANBUL' gibi adresleri tam olarak tespit edemediğimi gördüm. Bu nedenle, `([A-ZĞÜİÖÇŞ0-9][\.:\/]*[\t]*)` sorgusunu genişleterek `((0-9\.:\/A-ZĞÜİÖÇŞ)*[\t]*)((A-ZĞÜİÖÇŞ0-9)|:)(0-9\.:\/a-zA-ZĞÜİÖÇŞğüöçş)*[\t]*)` şekline dönüştürdüm. Sorgu bu haliyle aşağıdaki haline ulaştı.

`(T.*NO:.*)((0-9\.:\/A-ZĞÜİÖÇŞ)*[\t]*)((A-ZĞÜİÖÇŞ0-9)|:)(0-9\.:\/a-zA-ZĞÜİÖÇŞğüöçş)*[\t]*)+([/]\s*[İi](stanbul|STANBUL))`

Bu aşamadan sonra, 'Kalköy Deniz Otobüsü İskelesi önü Rasimpaşa Kadıköy/İstanbul' gibi adreslerin aslında tespit edilemediğini fark ettim. Bu tür uç durumlar için 'önü' ifadesini regex'ime ekledim. Sorgu, bu değişiklik ile aşağıdaki şekle geldi.

`(T.*NO:.*)((0-9\.:\/A-ZĞÜİÖÇŞ)*[\t]*)((A-ZĞÜİÖÇŞ0-9)|önü|:)(0-9\.:\/a-zA-ZĞÜİÖÇŞğüöçş)*[\t]*)+([/]\s*[İi](stanbul|STANBUL))`

Bu şekliyle, sorgunun tespit edemediği bir adres kalmadı. Bu durum beni oldukça memnun etti. Regex'in son halini, okunabilirliği artırmak amacıyla renklendirdim.

Sorgunun Son Hali:

`(T.*NO:.*)((0-9\.:\/A-ZĞÜİÖÇŞ)*[\t]*)((A-ZĞÜİÖÇŞ0-9)|önü|:)(0-9\.:\/a-zA-ZĞÜİÖÇŞğüöçş)*[\t]*)+([/]\s*[İi](stanbul|STANBUL))`

<https://regex101.com> üzerinde ECMAScript (JavaScript) ile çalıştırılan sorgunun sonucu aşağıda gösterilmiştir. Ayrıca, Python ile de çalıştırarak aynı sonuçları elde ettim. Sonuç görselini aşağıya ekledim. 😊😊😊 Adreslerin **%100'ünü** buldum 🐼🐼🐼

! / (T.*N0:.*)([0-9\.\-\/A-ZGÜİÖÇŞ]*[\t.*])([A-ZGÜİÖÇŞ0-9][önü]:)[0-9\.\-\/a-zA-ZGÜİÖÇŞgüröçş]*[\t.*)* / gm
([\/]\s*[!i](stanbul|İSTANBUL))

TEST STRING

Aşağıda adres bilgileri verilmiş olan farklı mekanlara farklı gruplar tamirat-tadilat, temizlik ve nakliye için gidecektir.␣
Ahmet Bey ve ekibi tamirat ve tadilat için Hürriyet Mah. Abdi İpekçi Cad. No:38/C KARTAL/İSTANBUL adresine 1 Aralık 2021 tarihinde 10 gün süre ile gidecektir.␣
Bu adresteki işler tamamlandıktan sonra Yıldırım Mah. Kemal Paşa Sok. No:8 BAYRAMPAŞA/İSTANBUL ve üç gün sonra da Top Ağacı M. Dereboyu Caddesi No:47 ÜMRANIYE,␣
Cumhuriyet Mahallesi Dr. Sadık Ahmet C. No:51/B KÜÇÜKÇEKMECE / İSTANBUL adreslerine gidelecektir.␣
Alper Bey'in sorumluluğunda olan temizlik ekipleri beş ayrı gruba ayrılarak 5 Aralık 2021 tarihinde aşağıdaki 5 farklı adrese temizlik için gönderileceklerdir.␣
Karaköy Deniz Otobüsü İskelesi önü Rasimpaşa Kadıköy/İstanbul␣
Yeni Mah. Dr. Ömer Besimpaşa Cad. No:64 Beykoz / İstanbul␣
Göztepe Mahallesi Tütüncü Mehmet Efendi Caddesi Rıdvanpaşa Sokak Kadıköy/İstanbul␣
Bağdat Cad. No:445 Kadıköy/İstanbul␣
PTT Evleri Bahçeköy Cad. No:53 Sarıyer/İstanbul␣
Mahmut Bey'in Lojistik firması da İstanbul'un farklı semtlerine beyaz eşya taşıyacaktır. Taşıma işinin 1-15 Aralık tarihleri arasında tamamlanması gerekmektedir.␣
Adresler öncelik sırasına göre aşağıda verilmiştir.␣

YENİBOSNA METRO İSTASYONU BAKIRKÖY / İSTANBUL␣
KENNEDY CAD. SIRKEÇİ ARABALI VAPUR İSKELESİ FATİH / İSTANBUL␣
YAVUZTÜRK MAH. KARADENİZ CAD. NO:2 ÜSKÜDAR / İSTANBUL␣
HASANPAŞA MAH. FAHRETTİN KERİM GÖKAY CAD. KADIKÖY / İSTANBUL␣
KARAKÖY YER ALTI GEÇİDİ NO:24 BEYOĞLU / İSTANBUL␣
ÖRNEK MAH. DOĞ. ARS. BLV. FIKRİ SÖN CAD. A GİRİŞİ NO:215/9/2 ESENYURT / İSTANBUL␣
GÜRSEL MAH. 28 NİSAN CAD. NO:4/B KAĞITHANE / İSTANBUL␣
ATATÜRK MAH. ALEMDAĞ CAD. NO:61 ÜMRANIYE / İSTANBUL␣
YILDIZ POSTA CAD. TÜRK TELEKOM ÖNÜ GAZETE BAYİİ BEŞİKTAŞ / İSTANBUL␣
ORTAÇEŞME SONDURAK BEYKOZ / İSTANBUL␣
MEHMET AKİF ERSOY MAH. NATO YOLU BOSNA BULVARI NO: 115 ÜSKÜDAR / İSTANBUL␣
DERBENT MAH. DEREİÇİ OTOBÜS SON DURAK SARIYER / İSTANBUL␣
LEVAZİM MAH. KORU SOK NO:7A BEŞİKTAŞ / İSTANBUL␣
19 MAYIS MAH. ŞEMSETTİN GÜNALTAŞ CAD. NO:144 KADIKÖY / İSTANBUL␣
YUNUS EMRE MAH. VEYSEL KARANI CAD. NO:105 SANCAKTEPE / İSTANBUL␣
ÇAMLIK MAH. MUHSİN YAZICIOĞLU CAD. NO:6/B ÇEKMEKÖY / İSTANBUL␣
AŞAĞI DUDULLU MAH. ALEMDAĞ CAD. NO:465/A ÜMRANIYE / İSTANBUL␣
DENİZ KÖŞKLER MAH. ESKİ EDİRNE ASFALTI NUR APT. NO:2 AVCILAR / İSTANBUL␣
KUYUMCU KENT ATÖLYE DURAĞI 29 EKİM CAD. LADİN SOK. K:1 NO:6 BAHÇELİEVLER / İSTANBUL␣
FEVZİ ÇAKMAK MAH. AHMET KOCABİYİK SOK. ANADOLU APT. NO: 3-A KÜÇÜKÇEKMECE / İSTANBUL␣
KARTALTEPE MAH. DİLEK SOK. NO:20 SEFAKÖY METROBÜS ALTGEÇİT KÜÇÜKÇEKMECE / İSTANBUL␣
KIRAZLIDERE MAH. MARAŞAL FEVZİ ÇAKMAK CAD. OYA SOK NO:2/1 ÇEKMEKÖY / İSTANBUL␣
KARTAL KÖPRÜSÜ ALTI METRO ÇIKIŞI ANKARA İSTİKAMETİ ES YAN YOL KARTAL / İSTANBUL␣
GÜRPINAR MAH. MİLLET CAD OKUTAN İŞ MERKEZİ NO:1 BEYLİKDÜZÜ / İSTANBUL␣
FEYZULLAH MAHALLESİ ŞEHİT HİKMET ALP CAD ADATEPE SİTESİ B1 BLOK DÜKKAN 2 MALTEPE / İSTANBUL␣
HIRKA-I SERİF MAH. MAH. LÜLECİYEKTA CAD. LÜLECİ YEKTA SOK. NO: 1 DAİRE: 1 KAT: 1 FATİH / İSTANBUL␣