



T.C.
SAKARYA ÜNİVERSİTESİ

BİLGİSAYAR VE BİLİŞİM BİLİMLERİ FAKÜLTESİ
BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ
PROGRAMLAMA DİLLERİNİN PRENSİPLERİ ÖDEV RAPORU

Dosya İşlemleri ve Düzenli İfadeler

Hazırlayan:

G181210023 - Muhammed Melih Çelik

SAKARYA

Mart, 2020

Programlama Dillerinin Prensipleri Dersi

Dosya İşlemleri ve Düzenli İfadeler

Muhammed Melih Çelik

G181210023 – 2/C

Özet

Herhangi bir uzunlukta ve sayfalarca içerikte olabilecek bir dosyayı okuyup içerisinde bulunan kelimelerin, sesli harflerin, cümlelerin, eposta adreslerinin, ve web adreslerinin kaç adet olduğu problemini çözmek için Java programlama dilinde düzenli ifadeler ve dosya işlemlerinden yararlandım. Java içerisindeki dosya okuma sınıflarını kullanarak icerik.txt isimli dosyayı programıma okuttum ve ilgili düzenli ifadeleri ilgili fonksiyonlar içerisinde kullanılmasıyla hesaplama işlemi tamamladım. Bu işlemleri gerçekleştirirken özellikle mail ve web site içerisindeki noktaların cümle saymada dikkate alınmamasına ve mail ile web sitelerinin karışmamasına özen gösterdim. Ödevi hazırlarken yayınlanan duyurular doğrultusunda birkaç değişikliğe gittim bunlar web adres sayısının kontrolü ile ilgiliydi. Elde ettiğim verileri kullanıcıyı bilgilendirmek amacıyla istenilen listeleme formatında ekrana bastırdım.

1. Geliştirilen Yazılım

Ödev içeriği dökümanında istenilen işlemleri gerçekleştirmek için statik fonksiyonlar tanımladım , çünkü istediğim işlemlerin bir kez gerçekleştirilmesi yeterliydi ve sınıftan obje üretme ihtiyacım yoktu.

Dosyayı okuyacak bir String tipi statik fonksiyon , kelime Dizisini elde edeceğim String dizi tipinde bir statik fonksiyon , sesli harf sayısı, kelime sayısı, cümle sayısı, mail sayısı ve web site sayısını bulacağım int tipinde statik fonksiyonlar tanımladım.

String tipi dosyaOku() fonksiyonumda Scanner sınıfını kullanarak dosya sonuna kadar icerik.txt'yi okuttum ve String bir değişkene atadım en sonunda bu değeri döndürdüm. Geri dönen bu String değerini kelimeDizisi() isimli fonksiyona parametre olarak gönderdim ve bu fonksiyon içerisinde String'i split ederek bir diziye atadım ve en son bu diziyi geri döndürdüm.

Sesli harf sayısını bulmak için String ifadenin kendisini kullandım Pattern ve Matcher sınıflarını kullanarak sesli harf sayısını bulduracak regex'i parametre olarak gönderdim. Parametre olarak gönderilen icerik Stringi içerisinde ifademle eşleşen sesli harfler kadar sesli harf sayısını artırdım. Sesli harfleri bulurken kullandığım Regex ifadesi: "[AEİİÖÜÜaeıioöüü]" 'idi böylelikle tüm büyük ve küçük sesli harfleri taramış oldu.

Kelime sayısını bulmak için kelime dizisinin uzunluğundan yararlandım ve kelimeSayisi() fonksiyonunda elde ettiğim tam sayı tipindeki uzunluğun değerini geri döndürdüm.

Cümle sayısını bulmak için yine bir regex ifadesi ve String dizimi kullandım. String dizimin her elemanı için gönderdiğim ifade eşleşiyor mu diye kontrol eden yapımda kullandığım Regex ifadesi: "\\.\$" idi böylelikle sonu nokta ile biten ifadeler cümle olarak kabul edilmiş oldu yani tamda dökümanda belirtildiği gibi.

Mail sayısını bulmak için yine mail'e özel bir regex ifadesi ve String dizimi kullandım ve yine dizinin her elemanı ile regexim eşleşiyor mu diye baktım. Mail sayısının bulmak için kullandığım Regex ifadesi: "^[_A-Za-z0-9-\\+]+(\\.[_A-Za-z0-9-\\+]*@[A-Za-z0-9-\\+](\\.[A-Za-z0-9-\\+])*(\\.[A-Za-z]{2,})\$)" idi böylelikle içerisinde @ işareti bulunan büyük küçük harfleri sayıları ve çizgi gibi karakterleri de göz önünde bulunduran bir yapı kurulmuş oldu. Noktaları kullanırken kaçış karakterlerini kullanmaya ve hata yapmamaya özen gösterdim. Mail ifadesindeki en sondaki noktadan sonraki ifadenin minimum değerini 1 yaptım ki cümle ile karışmasın yani noktadan sonra en az herhangi bir karakterin olmasını beklesin çünkü mail içerisindeki noktaların cümle saymada dikkate alınmaması gerekiyordu. Mail sayısını bulmayı da bu şekilde sağladım.

Site sayısı içinde benzer bir yapı kullandım bir regex parametresi ve yine String dizim parametre olarak alındı. Site sayısını bulurken kullandığım Regex ifadesi: "(http:\\\\www\\.|https:\\\\www\\.|http:\\\\|https:\\\\)?[A-Za-z0-9]+([\\-\\.]{1}[A-Za-z0-9]+)*\\. (edu\\.tr\$|edu\$|com\\.tr\$|com\$|net\\.tr\$|net\$|org\\.tr\$|org\$)" idi. İnternet sitelerinin başında http,https,http://,https:// gibi ifadeler bulunabilirde bulunmayabilirde bunu ? işareti ile sağladım. İnternet sitelerinin alan adlarında a'dan z'ye büyüklü küçükü harfler ve rakamlar bulunabilir ve bilgi sistemindeki duyuruya göre alan adlarının sonu .com, .edu, .net, .org, .com.tr, .edu.tr, .net.tr, .org.tr ile bitmelidir. \$ işareti ile bunu sağlamış oldum veya | işaretini kullanarak bu opsiyonları gezdim.

Bu fonksiyonlardan dönen tamsayı değerlerini istatistik() fonksiyonuna parametre olarak gönderdim ve bilgiyi ekrana bastım.

Problemin çözümünü başta String bir ifadenin karakter dizilerini gezerek yapacaktım fakat rapordaki kısıtlamaları gerçeklemenin daha bir kolay yolu olan düzenli ifadeleri kullanmayı tercih ettim. Bu sayede hem daha hızlı ve doğru sonuç alırken hemde problemin çözümü üzerinde daha kolay değişikliğe gidebildim.

Dosya okuma işlemi için ise Scanner sınıfını kullanmayı tercih ettim. Başta performans açısından BufferedReader'ı kullanmayı tercih etmeyi düşünsemde derste Scanner gördüğümüz için Scanner sınıfını kullandım.

Ödevde sesli harfleri bulmak için Türkçe dilini baz aldım fakat mail sayısı ve web sitelerinde böyle bir şey belirtilmediği için ve normalde de “ğüşöçİĞÜŞÖÇ” gibi harfler bulunmadığı için Türkçe karakterleri mailler ve web siteleri için almadım ve raporda kelimeleri bir boşluk ayıracaktır dediği için String ifadeyi .split(“ ”) ile kelimelere ayırdım bir boşluk ifadesinden dolayı herhangi başka bir kontrol mekanizması kurma ihtiyacı hissetmedim.

Ödevin verilme amacının Java’da dosya işlemlerinin kavranması ve programlamada düzenli ifadelerin işimizi nasıl kolaylaştırabileceğinin kavratılması ve araştırarak düzenli ifadelerin öğrenilmeye sevk edilmesi için verildiğini düşünüyorum.

ÇIKTILAR VE SONUÇLAR

Raporda verilen icerik.txt’yi okuttuğum zaman , icerik.txt:

fatihadak@sakarya.edu.tr sakarya.edu.tr www.google.com Esentepe Kampüs sakarya.com iletisim@ford

Elde ettiğim sonuçlar şu şekilde oldu:

Toplam Sesli Harf Sayısı : 33
 Toplam Kelime Sayısı : 7
 Toplam Cümle Sayısı : 0
 Toplam Mail Sayısı : 1
 Toplam Web Sitesi Sayısı : 3

Birde şu şekilde bir txt okuttum:

fatihadak@sakarya.edu.tr sakarya.edu.tr www.google.com Esentepe Kampüs sakarya.com iletisim@ford
 cumledenemesi. iletisim@ford.com.tr merhaba sitemibu.xyz

Elde ettiğim sonuçlar şu şekilde oldu:

Toplam Sesli Harf Sayısı : 52
 Toplam Kelime Sayısı : 11
 Toplam Cümle Sayısı : 1
 Toplam Mail Sayısı : 2
 Toplam Web Sitesi Sayısı : 3

Ödev dökümanı ve bilgi sistemindeki duyurular ışığında bir dosyayı okuyan toplam sesli harf , kelime , cümle , mail ve web sitesi sayısını bulan bir Java programı hazırladım.

© 2020 Sakarya Üniversitesi.

Bu rapor benim özgün çalışmamdır. Faydalanmış olduğum kaynakları içerisinde belirttim. Her hangi bir kopya işleminde sorumluluk bana aittir.

Anahtar Kelimeler: dosya okuma, düzenli ifadeler, split fonksiyonu
