

ONDOKUZ MAYIS ÜNİVERSİTESİ
BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ
PROGRAMLAMAYA GİRİŞ – I DERSİ
2020 – 2021 GÜZ DÖNEMİ ÖDEV – 6

Son Teslim Tarihi: 30.12.2020 Saat 23.59

Geç gönderilen ödevler değerlendirmeye alınmayacaktır.

Aşağıda verilen sorulara ilişkin çözümleri Ruby programlama dilini kullanarak kodlayınız. Cevapları, ders kapsamında önerdiğimiz (VS Code, Sublime, Atom) editörlerden herhangi biri üzerinde yazınız. Dosyalar isim formatı “ soru_no.rb ” şeklinde olmalıdır. Oluşturmuş olduğunuz her cevabı ise Türkçe karakterleri kullanmadan “ogrencino_ad_soyad_odev6” şeklinde bir klasöre ekleyiniz. “123_durmus_sahin_odev6” gibi. Son olarak oluşturmuş olduğunuz klasörü sıkıştırılmış dosya şeklinde (.tar, .zip, .rar gibi) “ogrencino_ad_soyad_odev6.(tar/zip/rar)” şeklinde Google Classroom’a yükleyiniz. Ek olarak paylaşılacak olan txt dosyasını ödev klasörünün içine eklemeyi unutmayınız.

NOT: Ödevlerde kod düzenine dikkat ediniz. Puanlamaya etki edecektir. Ayrıca ödevlerinizi lütfen kendiniz yapınız. Başkasından veya bir yerlerden bulduğunuz ödevler 0 puandır. Yukarıdaki talimatlara uymadığınız takdirde veya programınızın çalışmaması durumunda bu ödevden 0 puan alacaksınız.

Soru -1

Birkaç ayrıntıyı tartışmak istiyorum
Kendime ait bir odam olmasını istiyorum
Elimde olsa burada seninle yaşarım
İçeride kalmaktansa dışarı çıkmayı tercih ederim
Bugün okula gitmeyi tercih etmiyorum
Bana yardım edersen çok minnettar olurum
İki odalı bir daire kiralamak isterim
Bahçede bana yardımcı olmanı rica ediyorum
Burada ne olduğunu bilmek istiyorum
Duruma yardım ettiğini gerçekten sanmıyorum

a) Yukarıdaki cümlelerden oluşan ve sizinle paylaşılmış olan Cümleler.txt dosyası içerisindeki cümleleri oluşturan kelimeleri kullanarak bir hash (sözlük) yapısı oluşturmanız istenmektedir (Cümleler txt dosyasından okunacaktır). Örnek bir yapı şu şekilde olacaktır:

{"isterim":1, "yardım":2, "daire":3, "içerde":4 ..., "istiyorum":10, "birkaç":11, ...} gibi txt dosyasındaki bütün kelimeleri içeren bir yapı oluşturulacaktır (50 puan)

NOT: Bu adım için cümledeki bütün kelimelerin küçük harflere dönüştürülmesi gerekmektedir.

b) Bu yapı oluşturulduktan sonra sözlüğün sonuna OOV (Out of Vocabulary) yani cümleler de bulunmayan kelimeleri temsil etmek amaçlı bir anahtar-değer çifti eklenecektir. Örneğin sözlüğümüz içinde 10 tane anahtar-değer çifti olsun. Bu durumda 11. çift olarak yukarıdaki kurduğunuz sözlük yapısına uygun şekilde "OOV" ve 11 eklenecektir. (15 puan)

c) Cümleler.txt dosyasındaki bütün cümleler, oluşturulan sözlük kullanılarak vektörleştirilecektir (vektör ile dizi aynı şeylerdir). Örneğin txt dosyasındaki ilk iki cümlelerin vektörleştirilmiş hali şu şekilde olacaktır (15 puan).

Birkaç ayrıntıyı tartışmak istiyorum -> [11, 8, 42, 10]

Kendime ait bir odam olmasını istiyorum -> [33, 40, 18, 43, 26, 10]

NOT: Cümleler içerisindeki aynı kelimeleri sadece bir sayı temsil edecektir. Örneğin yukarıdaki cümlelerde "istiyorum" kelimesi iki cümlede de geçmektedir ve karşılık gelen değer 10'dur. Bunu yapmak için küme mantığını kullanabilirsiniz.

d) Kullanıcıdan girdi olarak alınan bir cümleye karşılık gelen vektörün çıktı verilmesi istenmektedir. Örneğin;

"Birkaç ayrıntıyı tartışmak istemiyorum" -> [11, 8, 42, [OOV]] (OOV yazan yere sözlüğünüzde ne karşılık geliyorsa o yazılacaktır). (20 puan)

!!! Bu dökümanda paylaşılan çıktılar ile sizin alacağınız çıktılar tamamen aynı olmayabilir. Bu dökümandakiler örnek temsil etmesi açısından paylaşılmıştır !!!

!!! Cevaplar.txt dosyası içerisindeki cümlelerin UTF-8 formatında okunması gerekebilir. Bu durumu göz önüne alınız !!!