



T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
BİLGİSAYAR MÜHENDİSLİĞİ

HESAPLAMALI ANLAM BİLİM
ÖDEV 1

Hazırlayanlar
Derya TANYILDIZ, 21501111
Aleyna ER, 21501031

Ders Sorumlusu
Doç.Dr. Mehmet Fatih AMASYALI

20.03.2022

İÇİNDEKİLER

1.GİRİŞ.....	1
I. Ödevin Amacı	2
II. Ödevin Aşamaları	3
III.Kelime Bulma Oyunu	4
IV.Cümle Oluşturma Uygulaması.....	8

1.1. Ödevin Amacı

Ödevin ilk aşamasında toplanan hayat bilgisi veri seti kullanılarak bu ilişkiler doğrultusunda cümleler üreten bir uygulama tasarlanması ve bu uygulamaya ek olarak kelimeler arasındaki ilişkiler kullanılarak yakınlık ve uzaklıkları kullanarak bir oyun tasarlanması.

1.2. Ödevin Aşamaları

Ödev kapsamında ilk olarak verilerin .pl dosyaları python pandas yardımı ile düzenlenerek istenilen formata getirildikten sonra csv dosyaları şekilde kaydedildi. Daha sonra cümle oluşturma kısmı için Türkçe'nin cümle yapıları göz önünde bulundurularak özne, nesne, sıfat, dolaylı tümleç, zarf tümleci, edat ve yüklemelerden oluşan ilişki listeleri hazırlandı. Daha sonra bu listelerin içerisinde cümle yapısına uygun sıralama ile rastgele kelimeler seçilerek cümleler oluşturuldu. Buna ek olarak yine hayat bilgisi veri seti kullanılarak benzerlik kurulup bir sıcak-soğuk oyunu tasarlandı.

1.3. Ödevin Çıktıları

Ödev kapsamında oluşturulan cümle üretme çalışmasında “Özne + Sıfat + Nesne + Dolaylı Tümleç + Yüklem” yapısı kullanılarak oluşturulan ilişki listeleri içerisinde cümleler oluşturuldu ve bir arayüz ile gösterimi sağlandı. Ancak kelimeler arasındaki ilişki sadece verilen veri seti ile kurulduğu için istenilen başarıyı gösteremedi ve rastgele kelimeler ile cümleler kuruldu. Daha da başarılı sonuçlar almak için ödev ek olarak bir de benzerlikler kullanılarak sıcak-soğuk oyunu eklendi.

2.1. Veri Ön İşleme

İsim ve fiiller için sağlanan veriler (.pl dosyaları) txt olarak okunmuştur. cleanText fonksiyonu yazılarak metindeki “ilişki” keywordü ve noktalama işaretleri çıkarılmış, tanımlanamayan karakterler ise yerine başka bir harf konarak değiştirilmiştir. Kullanıcılardan alınan ilişki değeri metinlerine, açık kaynaklı temin edilen Türkçe Spell Checker ([1])* kütüphanesi ile yazım denetimi uygulanmıştır. Ne yazıkki, bu kütüphaneden istenen verim alınamamıştır, örneğin “acıklamis” kelimesini “açıklamış” olarak düzeltmemiş ve olduğu gibi bırakmıştır. İlişkiyi belirten sorularda da yazım hataları bulunmaktadır, bu soruların doğru yazımı listede tutulmuş ve yazım yanlış olan ilişki soruları listede en çok benzeyen soruya dönüştürülmüştür. Sonrasında bu ilişki sorularından (türleri) feature’lar (sütunlar) oluşturulmuştur. Veri çoklamasını önlemek için groupby fonksiyonu kullanılmıştır, bir ilişki türü birden fazla değer aldığından (birçok insan tarafından etiketlendiğinden) bu değerler hücrelerde liste olarak tutulmuştur.

Kelimeyi Bul oyunu için, isim ve fiiller verisi birleştirilerek işlenmiştir. Bu tabloda sütunların yeri uzaklık belirttiğinden, önemli görülmesine göre isim&fiil ilişki türleri (sütunları) sıralanmış ve sütunların yeri manipüle edilmiştir. Oyunda hacim ve ağırlık ilişkisi kullanılmadığından bu sütunlar drop edilmiştir. Tablonun son hali 2.3. bölümünde bulunan Şekil 6’da gösterilmiştir.

2.2. Kelimeyi Bul Oyunu (hot&cold)

Kelimeyi Bul oyunu temel olarak bir sıcak-soğuk oyunudur. Kullanıcının girdiği deęerin istenilene anlamsal olarak yakın olmasına göre sıcak-soğuk dönüşü yapılır. Oyunun amacı, random seçilen kelimenin kullanıcı tarafından tahmin edilmesidir. Kullanıcı girdi olarak kelime veya kelime grubu girebilmektedir.

Oyun için veri seti düzenlenirken farklı bir yaklaşım gözetilmiştir. İsim ve fiiller dosyası birleştirilerek kullanılmış, sütunların yeri manipüle edilmiştir. Sütunlar, kelime, kelimenin tanımı daha sonrasında ise bir isim bir fiil sorusu (özellik) olmak üzere devam etmektedir. Bu sorular sıralanırken, bir kavramı en iyi ifade edebilecek özellikler göz önünde bulunmuş ve önemlerine göre en iyi özellik (soru) başa alınmıştır.

Random kelime seçildikten sonra, tabloda o kelimeye yakın kavramlar gezilir ve bir sözlükte (dictionary) uzaklık bilgileriyle beraber tutulur. Bu uzaklık bilgileri, ceza olarak da isimlendirilebilir, random kelimeye olan sütunsal ve satırsal uzaklık ile ölçülür.

	ilişki1	ilişki2	ilişki3	ilişki4	ilişki5	ilişki6
kelime1	13	14	15	16	17	18	
kelime2	1	2	3	4	5	6	
kelime3							
kelime4	7	8	9	10	11	12	
kelime5							
kelime6							
.							
.							
.							

Şekil 1. Örnek

Örneğin random seçilen kelime2 ise, anlamsal yakınlar sözlüğü oluşturulurken;

1: kelime2' nin sütunlarını gez, 1,3 gibi tek sayılı sütunlara geçtiğinde cezayı 1 arttır (isim sütunu,fiil sütunu ... olarak listelendiği için isim ve fiil sorularında eşitlik sağlanması için böyle yapılmıştır)

Sözlük = {1:1, 2:1, 3:2, 4:2 .. sütunDeğeri:ceza}

Sütun değerlerinde kelimeler ile ilişkili kavramlar bulunmaktadır ve bu kavramlar liste olarak tutulmaktadır. Öyle bir ilişki değeri yoksa boş liste döndürülür. Örneğin anlaşılabilir olması açısından, her ilişkide tek bir değer (kavram) varmış gibi anlatılmıştır.

2: kelime2 sütunları gezilip temel sözlük oluşturulduktan sonra, bu sözlükteki kavramların tabloda kelime olarak tutulup tutulmadığı kontrol edilmiştir. Örnekte, 1 numaralı değer, tabloda kelime4 olarak tutulduğu bulunmuş, daha sonrasında buradaki değerler uzaklıklarıyla beraber sözlüğe eklenmiştir. Başka bir kayda erişmenin cezası 5 puandır, buradaki değerlerin uzaklığı parent kelimenin cezası+5 olarak hesaplanmıştır. Sözlük = {1:1, 2:1, 3:2, 4:2, 5:3,6:3, 7:6, 8:6, 9:7, 10:7, 11:8, 12:8}

Başlangıç sözlüğünde tutulan değerlerin hepsi sözlükte kelime olarak var olmayabilir, böyle durumlarda pas geçilir. Örnekte, 4 numaralı değer tabloda kelime1 olarak tutulmaktadır, bir önceki işlemin aynısı uygulanır.

Sözlük = {1:1, 2:1, 3:2, 4:2, 5:3,6:3, 7:6, 8:6, 9:7, 10:7, 11:8, 12:8, 13:7, 14:7, 15:8,}

Sözlüğe eklenen her kavram için bu işlem uygulanır, ancak bu uygulamanın sınırları bulunmaktadır. Sözlükte bulunan bir kavramın (kelimenin) cezası (tahmini yapılacak kelimeye uzaklığı) en fazla 20 olabilir. Bu şart sağlanmayana kadar işlem devam ettirilir.

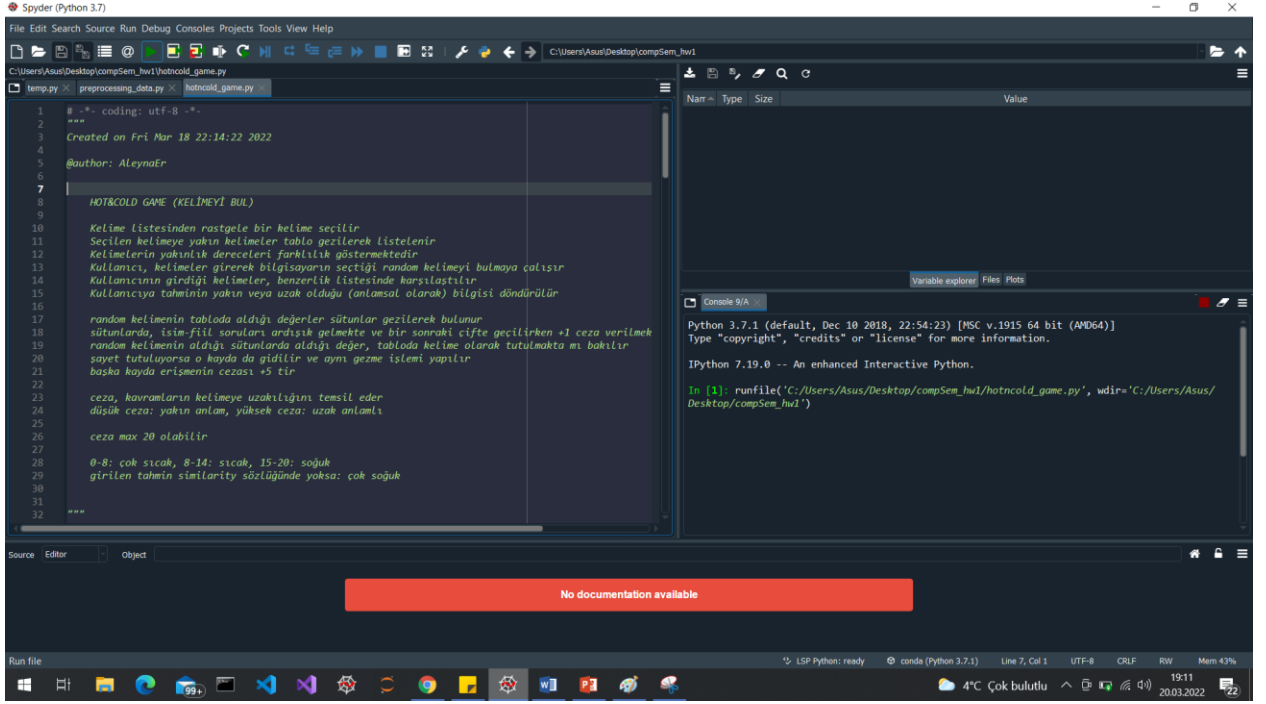
Anlamsal yakınlar sözlüğü oluşturulduktan sonra, kullanıcıdan input alınarak random seçilen kelimeyi tahmin etmesi istenir. Eğer kullanıcının girdiği kelime (yahut kelime grubu) sözlükte bulunuyor ise uzaklık değerine göre kullanıcıya bilgi döndürülür;

ceza (uzaklık) değeri; 0-8: çok sıcak, 8-14: sıcak, 15-20: soğuk dönütü yapılır.

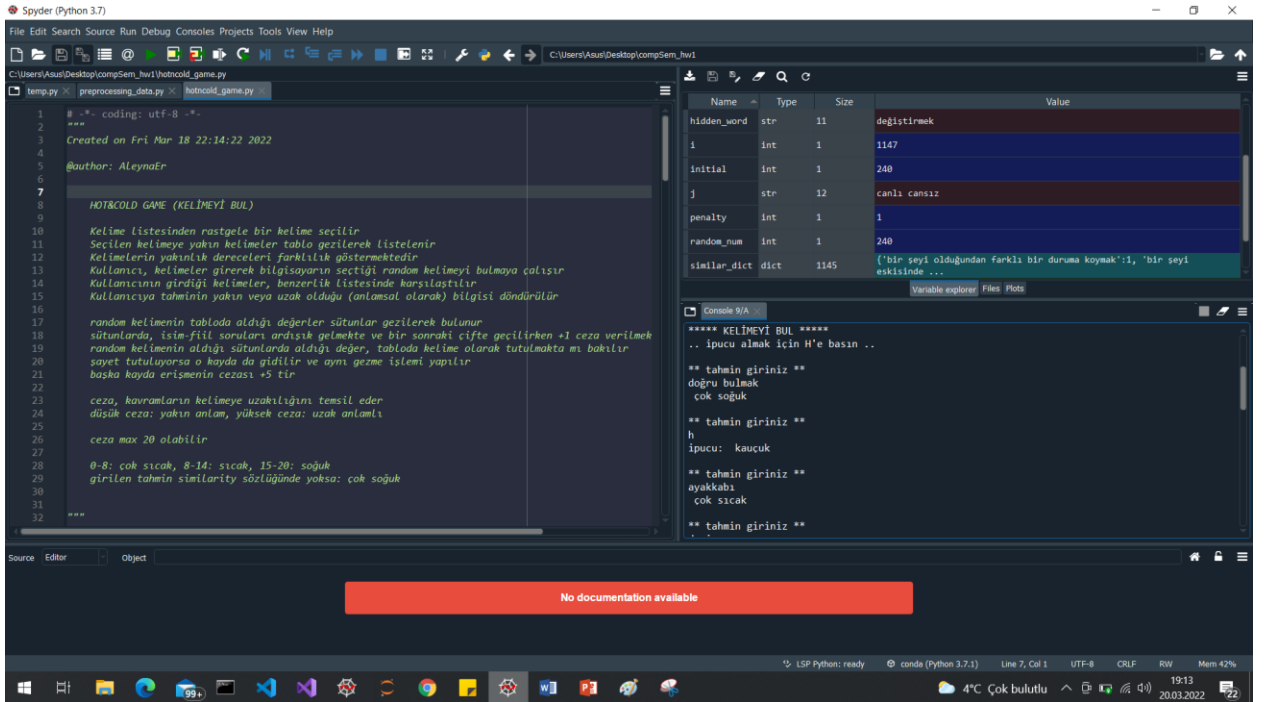
Dönütün “çok sıcak” olması girilen kavramın istenilene (random seçilen gizli kelime) anlamsal olarak çok yakın/benzer olduğunu gösterir. Dönüt “çok soğuk” ise girilen değer anlamsal yakınlar sözlüğünde bulunmamaktadır.

Oyunda, kullanıcının 5 tahmin ve 3 ipucu isteme hakkı bulunmaktadır. Oyun her tekrar başlatıldığında, yeni bir random kelime seçilir ve o kelimenin anlamsal yakınları hesaplanarak sözlük oluşturulur.

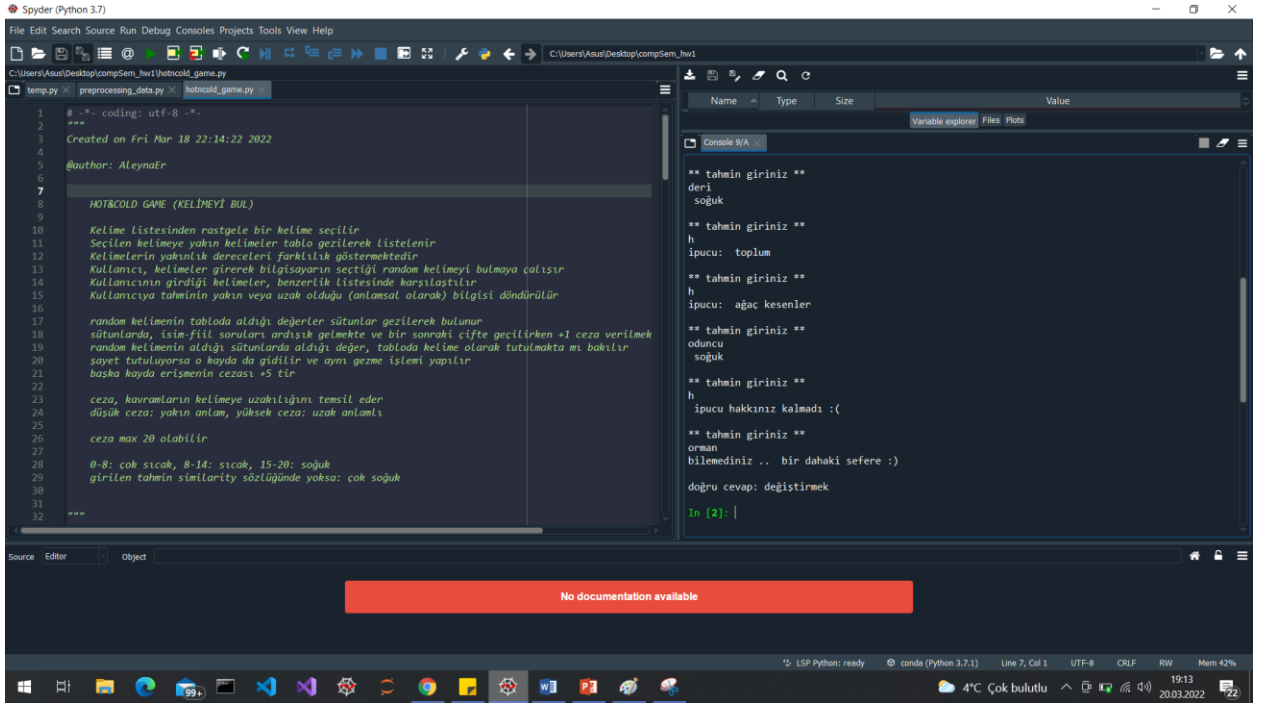
Oyundan kesitleri aşağıda verilmiştir:



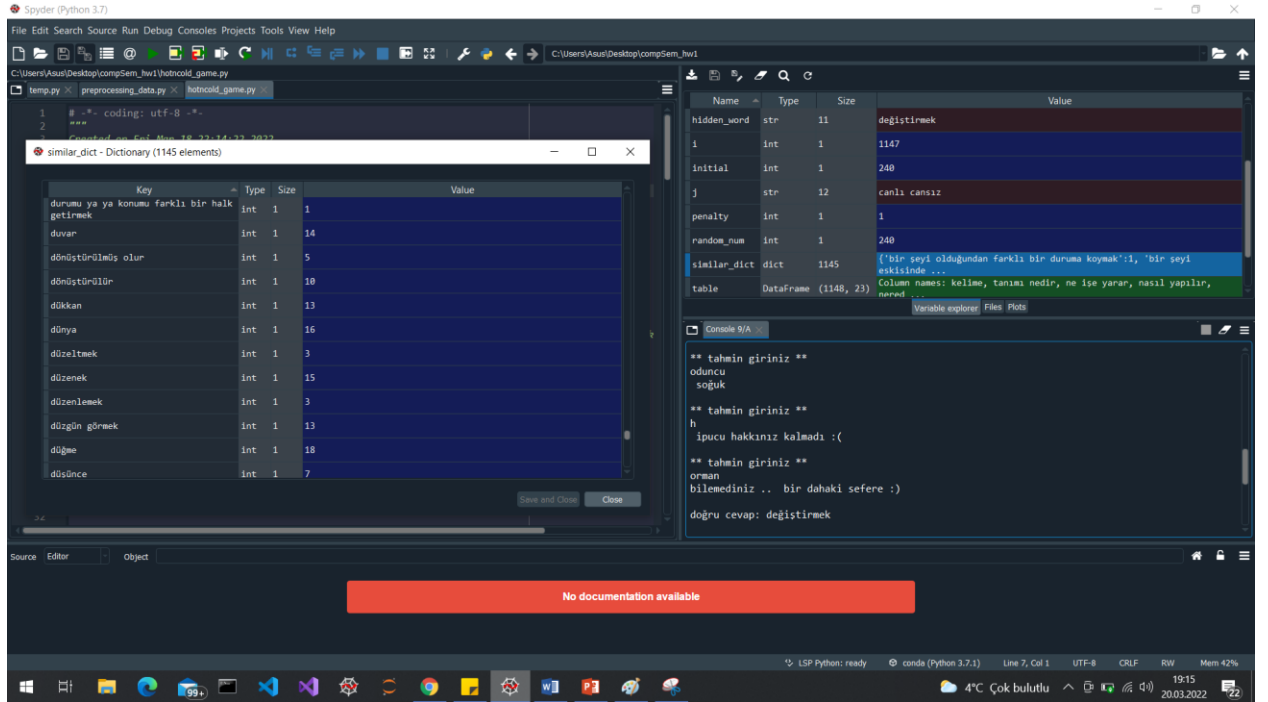
Şekil 2. Oyun başlatıldığında arka planda yakın anlamlılar sözlüğü oluşturulur



Şekil 3. Yakın anlamlılar sözlüğü oluşturulduktan sonra kullanıcıdan tahmin girdileri alınır



Şekil 4. Tahmin girdileri dönütü ve ipucu verme, doğru cevabın açıklanışı



Şekil 5. Doğru cevaba ait yakın anlamlılar sözlüğü, kavramlar ve cezaları (uzaklıkları)

Sütunların önem sırası ise aşağıdaki gibidir:

'tanımı nedir', '>'ne işe yarar' 'nasıl yapılır', '>'nerede bulunur' 'niçin yapılır', '>

'yanında neler bulunur' , 'ne olunca yapılır' > 'içinde neler bulunur', 'yapınca ne olur' >

'hammaddesi nedir', 'neyi kimi yapılır' > 'üst kavramı nedir', 'kim ne ile yapılır' >

'kim kullanır', 'neye kime yapılır' > 'sıfatları', 'nerede yapılır' > 'rengi',

'kim ne yapar' > 'şekli nasıl', 'fiziksel zihinsel' > 'canlı cansız'

Index	kelime	tanımı nedir	ne işe yarar	nasıl yapılır	nerede bulunur
0	acele etmek	['bir işi hızla yapmak', 'çabuk davranmak']	['']	['kösarak', 'koymak', 'kösamak']	['']
1	açıklamak	['yemek yeme gereksinimini duymak']	['']	['konusmak']	['']
2	açılmak	['açma işine konu olmak']	['']	['konusmak']	['']
3	acmak	['bir şeyi kapalı durumdan açık duruma getirmek']	['']	['itermek']	['']
4	affetmek	['yapılan hatayı hoşgörmek', 'hatayı hoşgörmek']	['']	['konuşarak', 'bağışlamak', 'hoşgörü', 'özür d...']	['']
5	akbil	['toplu ulaşımında kullanılan elektronik ücret t...']	['toplu taşımaya binmek', 'otobüse binmek', 's...']	['']	['vapur', 'büğrenci', 'otobüs', 'gişe', 'çanta'...
6	akıl etmek	['önmek veya çareyi zamanında düşünmek']	['']	['dusunerek']	['']
7	akordeon	['müzik yapmak için kullanılan körüklü çalgı']	['müzik', 'sarkı çalmak', 'eglençe', 'melodi ç...']	['']	['müzik dükkanları', 'mağaza', 'calgıcıları', ...']
8	akumulator	['araçlarda elektrik enerjisi kaynağı']	['devre çalıştırmak']	['']	['gemi']
9	akümülatör	['kimyasal enerji depolayan cihaz', 'elektrik ...']	['motora güç verir', 'enerji depolama', 'araba...']	['']	['gemi', 'otobüs', 'kamyon', 'taxis', 'hane', 'oto...']
10	akıl etmek	['', 'alınabilecek önlem iş isten geçmeden dü...']	['']	['akla gelin', 'çıkarma yaparak', 'düşünerek']	['']
11	alay etmek	['gölünç ve yönlerini küşseyerek eğlenme kon...']	['']	['çok', 'konuşarak', 'davranış', 'açık bulanok...']	['']
12	alci	['bir yapı malzemesi']	['tıra yapmak']	['']	['bina']
13	alışmak	['bir işi kolaylıkla yapabilmek']	['']	['tekrar etmek']	['']
14	almak	['bir şeyi veya kimseyi bulunduğu yerden ayırm...']	['']	['kolay', 'utanarak', 'dikkatlice', 'satın alt...']	['']
15	altın	['değerli bir elementi']	['elektrikli iletir']	['']	['yeryüzü']
16	altın	['']	['']	['']	['']
17	altın	['', 'taki yapısında kullanılan maden', 'ziyane...']	['iletin', 'ekonomi', 'süs esyası', 'aksesuar ...']	['']	['mücevherde', 'yer altı', 'altın madeni', ...']
18	alci	['sanatta', 'beyaz inşaat malzemesi', 'kaplama...']	['dekoratif', 'sekil vermek', 'kol kırıklarını...']	['']	['tedavi', 'hasta', 'dişi', 'medikal mağaza', ...']
19	alışmak	['', 'bir eylemi yapı yapı', 'intibak etmek', ...']	['']	['tekrarlayarak', 'eğitmek', 'çalışarak', 'zam...']	['']
20	ambalaj	['Ürünleri içinde barındıran plastik madde', ...']	['nesneleri korur', 'paketlemek', 'koruma sağl...']	['']	['mağaza', 'fabrika', 'esyanların dışında', 'ka...']
21	anahtar	['kilidin açılmasını sağlayan alet', 'dil', 'k...']	['açmak', 'kapıları açmak ve kilitlemek', 'kap...']	['']	['', 'kapının üstünde', 'kilit', 'çanta', 'evd...']
22	anlamak	['fikri kavramak', 'kavramak', 'olgu ya ya dav...']	['']	['okuyarak', 'kolay', 'detaylı', 'araştırmak', ...']	['']
23	anlatmak	['birinin bir konuyu ya ya şeyi anlamasını sağ...']	['']	['konuşarak', 'yazarak', 'ders', 'detaylı', 'g...']	['']
24	anmak	['bir objenin lafi geçmek', 'olayı', 'akla get...']	['']	['saygıyla', 'hatırlamak', 'saygı duruşu', 'ha...']	['']

Şekil 6. Kelimeyi Bul oyunu için kullanılan temizlenmiş tablodan kesit

2.3. Cümle Oluşturma (Sentence Generator)

Türkçe dili yapısına uygun olarak cümle yapısı oluşturuldu. Örnek vermek gerekirse;

“Özne + Sıfat + Nesne + Dolaylı Tümleç + Yüklem”

“Özne + Nesne + Yüklem”

kombinasyonları şeklinde isimler ve fiiller dosyaları kullanılarak “sentence generator” uygulaması geliştirildi.

Özneyi oluşturmak için isimler dosyasındaki “kelime” sütunu kullanıldı. Sıfatı oluşturmak için isimler dosyasındaki “sıfatları” ve “rengi” sütunları bir araya getirildi.

Nesne sorusuna cevap vermek için ise isimler dosyasından “kim kullanır?” ve fiiler dosyasından “kim/ne yapar?” soruları kullanıldı.

Fiileri oluştururken ise fiiler dosyasındaki “kelime” sütunu kullanıldı.

Bunlara ek olarak tümleçler oluşturuldu zarf tümleci oluştururken fiilere “nasıl yapılır” sorusu soruldu, dolaylı tümleçte ise isimlerden “nerede bulunur” ve fiilerden “nerede yapılır” ve “neye/kime yapılır” sorularına ait cevaplar kullanıldı.

Son olarak cümle yapısı olarak edatları bulmak için “kim/ne ile yapılır?” sorusuna cevap arandı.

Tüm bunlar değerlendirildiğinde hem dolaylı tümleç; edat ve zarf tümlecinden daha başarılı sonuç verdiği için o kullanıldı ve en basit cümle yapısı olan özne nesne ve yüklem ile cümleler oluşturuldu

Sentence Generator Application

Generate

yetkili kişiler istenilen renklerde hurda afet farzetmek

Şekil 7. Özne + Sıfat + Nesne + Dolaylı Tümleç + Yüklem ile oluşturulan cümle.

Sentence Generator Application

Generate

bilim insanları termometre ilan etmek

Şekil 8. Özne + Nesne +Yüklem ile oluşturulan cümle.

REFERANSLAR

- [1] <https://pypi.org/project/turkishnlp/>
- [2] https://github.com/aleynaer/hesaplamaliAnlambilim_hw1
- [3] <https://github.com/tanyildizderya/hesaplamalianlambilim>
- [4] <https://sites.google.com/view/mfatihamasyali/hesaplamal%C4%B1-anlambilim?authuser=0>