



GEOGRAPHY QUESTION ANSWERING

DOLU DOLU ANADOLU

A question answering (QA) model that can answer geographic questions
that are asked in Turkish free-text.

TUR

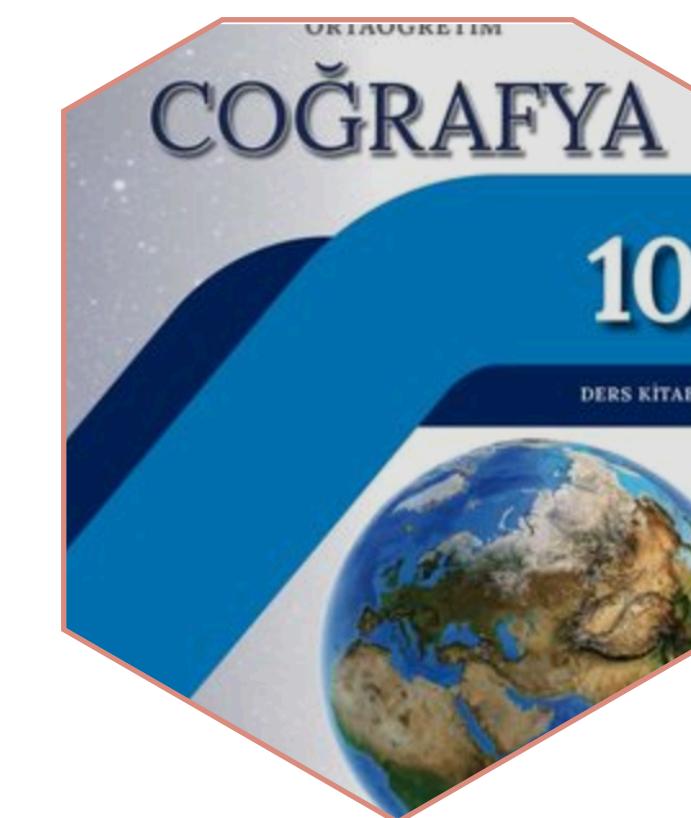
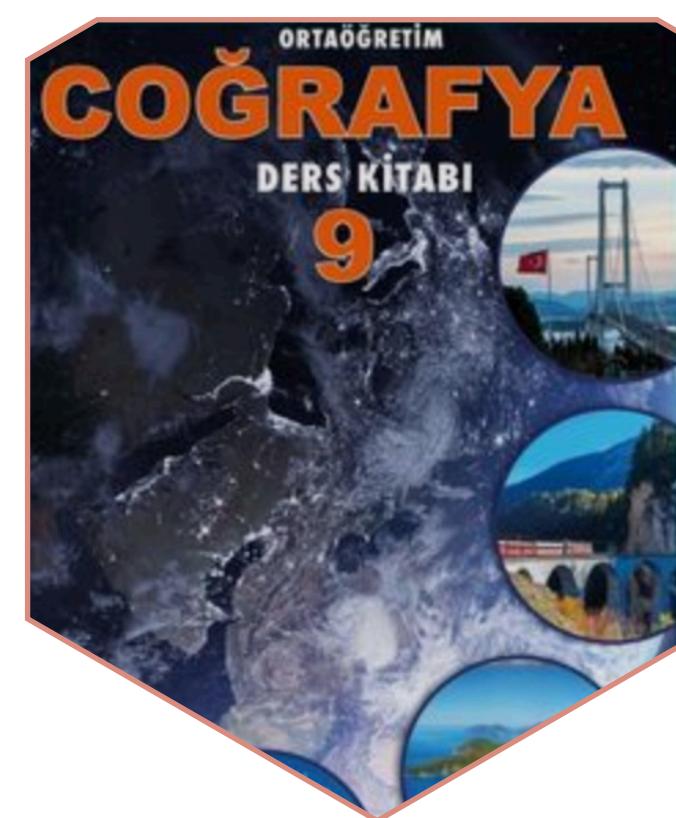
Dilara Gökay, Dilruba Reyyan Kılıç, Özge Dinçsoy

DATA SET:

High school geography books that are used by Turkish Education Ministry

QUESTIONS & ANSWERS:

Q&A pairs created by the students of Boğaziçi University, taking Cmpe493 course.



PREPROCESSING

Remove Punctuation

Remove Extra White Spaces

Solve Capital "i" Problem

Lemmatize

Lower Case

Create a pickle Dictionary

DICTIONARY

```
{  
  "226": {  
    "paragraph":  
      "Madenler, yaşamımızın birçok alanında çeşitli amaçlarla  
      işleyerek kullandığımız doğal kaynaklardır. İnşaat, otomotiv,  
      beyaz eşya, mutfak eşyaları, elektrikli ve elektronik eşya  
      üretimi gibi birçok alanda madenler kullanılmaktadır.",  
    "sentences": [  
      "Madenler, yaşamımızın birçok alanında çeşitli amaçlarla  
      işleyerek kullandığımız doğal kaynaklardır.",  
      "İnşaat, otomotiv, beyaz eşya, mutfak eşyaları, elektrikli ve  
      elektronik eşya üretimi gibi birçok alanda madenler  
      kullanılmaktadır."  
    ]  
  }  
}
```

encoding="utf-16" ~

partToSentences.pickle

METHODOLOGY

1

Find the Related Paragraph

Extract The Answer

2

$$tf\text{-}idf \text{ scoring}$$
$$w_{i,j} = tf_{i,j} \times \log \left(\frac{N}{df_i} \right)$$

$tf_{i,j}$ = number of occurrences of i in j

df_i = number of documents containing i

N = total number of documents

Reference:

Metin Madenciliği ile Soru Cevaplama Sistemi

Sevinç İlhan¹, Nevcihan Duru², Şenol Karagöz³, Merve Sağır⁴

¹Mühendislik Fakültesi Bilgisayar Mühendisliği Bölümü
Kocaeli Üniversitesi

METHODOLOGY

1

Find the Related Paragraph

2

Extract The Answer

a paragraph is returned

predict
sentence

tf-idf scoring

$$w_{i,j} = t f_{i,j} \times \log \left(\frac{N}{df_i} \right)$$

a sentence
from the
paragraph

Answer

METHODOLOGY

①

Find the Related Paragraph

Extract The Answer

②

Related Sentence

Question

Answer

EXAMPLE:

Extract The Answer

Almanya'da hangi iklim şartları hakimdir?

Related Paragraf: 3585

Almanya'da ılıman okyanusal ve karasal iklim şartları hâkimdir. Ülkenin kuzey ve kuzeybatı kesimlerinde batı rüzgârları ve Gulfstream sıcak su akıntısının etkisiyle ılıman okyanusal iklim görülür. Ülkenin güney ve orta kesimlerinde ise yükselti ve karasallığın etkisiyle karasal iklim koşulları egemendir. Ülke topraklarının yaklaşık üçte birini oluşturan ormanlar, güneydeki dağlık alanlarda geniş yer kaplar.

Related Sentence:

Almanya'da ılıman okyanusal ve karasal iklim şartları hâkimdir.

~~almanya da ılıman okyanusal ve karasal iklim şart hâkim~~

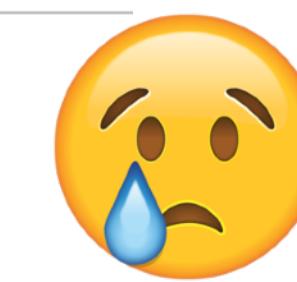
~~almanya da hangi iklim şart hakim~~



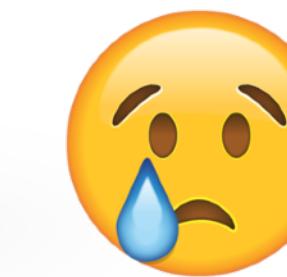
okyanusal ve karasal

EVALUATION

ITU Turkish Natural Language Processing Pipeline
ITU Türkçe Doğal Dil İşleme Yazılım Zinciri



 github.com/ahmetaa/zemberek-nlp



StanfordNLP

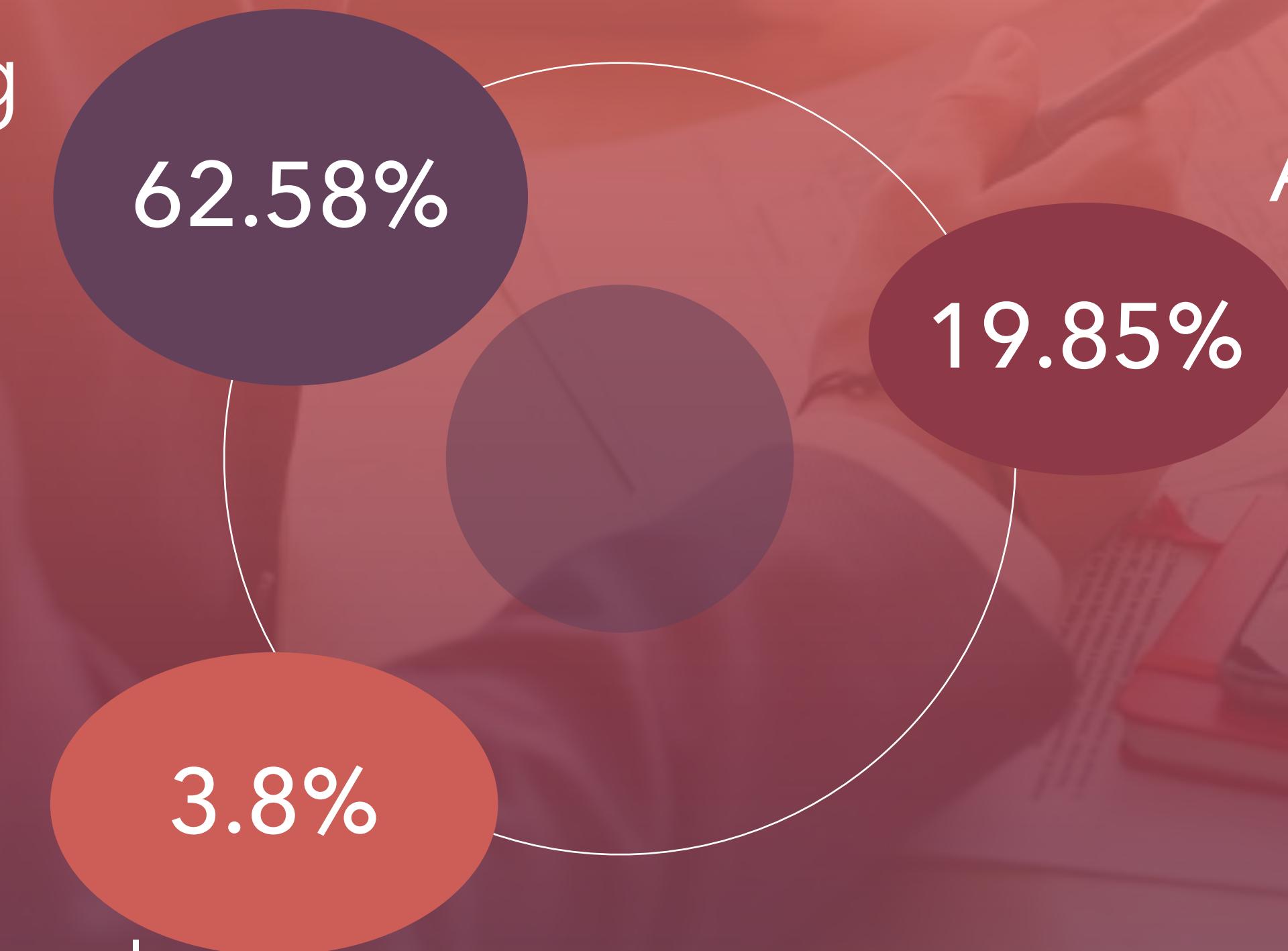
pypi v0.2.0 python 3.6 | 3.7



 github.com/akoksal/Turkish-Lemmatizer

RESULT

Paragraph
Finding



Answer Retrieval
(Accuracy)

Answer Retrieval
(Jaccard)



DOLU DOLU ANADOLU

DİLARA GÖKAY
DILRUBA REYYAN KILIÇ
ÖZGE DİNÇSOY