



**Bursa Teknik Üniversitesi**  
**Mühendislik ve doğa bilimleri Fakültesi**  
**Bilgisayar Mühendisi**  
**Doğal Dil İşlemeye Giriş**  
**(BLM0467)**

**Proje Raporu**

**Proje Adı:** Bir metnin okunabilirlik seviyesinin (ilk okul, lise, üniversite, teknik doküman) belirlenmesi

Video Linki : <https://youtu.be/5wpLlvdvqPg>

**Adı ve Soyadı:** Farhan Ahmad

**Okul Numarası:** 20360859096

## **Proje Başlığı:**

Metinlerin Okunabilirlik Seviyesinin Belirlenmesi

## **Amaç ve Hedefler:**

Bu projenin amacı, verilen metinlerin okunabilirlik seviyesini belirleyerek bu metinlerin hedef kitleye uygun olup olmadığını değerlendirmektir. Bu doğrultuda, Flesch-Kincaid okunabilirlik testini kullanarak metinlerin ilkökul, ortaokul, lise, üniversite veya teknik doküman seviyesinde olup olmadığını belirleyeceğiz.

## **Kapsam ve Sınırlar:**

Proje kapsamında, Türkçe yazılmış metinler üzerinde analiz yapılacaktır. Hece sayısının belirlenmesi için CMU Pronouncing Dictionary (CMUDict) kullanılacak, ancak sözlükte bulunmayan kelimeler için alternatif bir hece tahmin yöntemi uygulanacaktır. Bu proje, özellikle eğitim materyalleri, teknik dokümanlar ve çeşitli yazılı içeriklerin okunabilirlik seviyesinin belirlenmesi için faydalı olacaktır.

## **Problem Tanımı**

### **Problem Açıklaması:**

Okunabilirlik seviyesi, bir metnin ne kadar kolay anlaşılabilirliğini belirleyen önemli bir ölçüttür. Bu seviye, hedef kitleye uygun içerik oluşturmak için kritik bir öneme sahiptir. Özellikle eğitim materyalleri, teknik dokümanlar ve çeşitli medya içerikleri için metinlerin okunabilirlik seviyesinin doğru belirlenmesi gerekmektedir.

## Motivasyon:

Doğru okunabilirlik seviyesi belirlemek, eğitim materyallerinin, resmi belgelerin ve diğer yazılı içeriklerin etkinliğini artırır. Bu, okuyucuların metni daha iyi anlamalarını ve bilgilerden daha verimli yararlanmalarını sağlar. Ayrıca, eğitim materyallerinin hedef kitleye uygun hale getirilmesi, öğrenme sürecini destekler ve bilgi edinmeyi kolaylaştırır.

## Girdi ve Çıktı

- Girdi

Bir metin örneği:

**Burada örnek bir metin olacak. Metnin okunabilirlik seviyesini belirleyeceğiz.**

- Çıktı

Okunabilirlik seviyesi ve Flesch-Kincaid derece seviyesi:

**Metin: Burada örnek bir metin olacak. Metnin okunabilirlik seviyesini belirleyeceğiz.**

**Okunabilirlik Seviyesi: İlkokul (Flesch-Kincaid Grade Level: 3.67)**

## Yöntem ve Adımları

Bu çözüm, Flesch-Kincaid okunabilirlik testi kullanılarak uygulanmıştır. Flesch-Kincaid testi, metnin okunabilirliğini belirlemek için cümle başına düşen kelime sayısı ve kelime başına düşen hece sayısını kullanır.

## Adımlar

- Metni Bölme:
  - Metin, cümlelere bölünür.
  - Cümleler, kelimelere bölünür.
- Hece Sayısı Hesaplama:
  - Her kelimenin hece sayısı hesaplanır.
- Formül Uygulama:
  - Flesch-Kincaid okunabilirlik formülü kullanılarak okunabilirlik seviyesi hesaplanır:

$$\text{Grade Level: } 0,39 \left( \frac{\text{Toplam Kelime}}{\text{Toplam Cümle}} \right) + 11,8 \left( \frac{\text{Toplam Hece}}{\text{Toplam Kelime}} \right) - 15,59$$

- Seviye Belirleme:
  - Hesaplanan okunabilirlik derecesine göre metnin seviyesi belirlenir (ilk okul, orta okul, lise, üniversite, teknik doküman).

## Çözüm İçin Özgünlük

Bu çözüm, hece sayısını doğru bir şekilde belirlemek için nltk kütüphanesinin CMU Pronouncing Dictionary (CMUDict) veri tabanını kullanır. Bu, kelime başına düşen hece sayısını doğru bir şekilde hesaplamayı sağlar. Ayrıca, kütüphane dışı bir yöntem olarak, kelimenin hece sayısı tahmin edilemediğinde, sesli harflerin sayılması yöntemi de

kullanılmıştır. Bu, metnin okunabilirlik seviyesinin doğru bir şekilde belirlenmesini sağlar.

## Benzer Çözümler ve Kaynaklar:

- **Benzer Çözümler:**

- Flesch-Kincaid okunabilirlik testi yaygın olarak kullanılır ve birçok metin editörü ve eğitim yazılımında bulunur.
- Gunning Fog Index ve SMOG Index gibi diğer okunabilirlik testleri de benzer şekilde metin analizleri için kullanılabilir.

- **Kaynaklar:**

1- nltk Kütüphanesi: <https://www.nltk.org/>

2- CMU Pronouncing Dictionary (CMUDict):  
<http://www.speech.cs.cmu.edu/cgi-bin/cmudict>

3- Flesch-Kincaid Okunabilirlik Testi:  
[https://en.wikipedia.org/wiki/Flesch%E2%80%93Kincaid\\_readability\\_tests](https://en.wikipedia.org/wiki/Flesch%E2%80%93Kincaid_readability_tests)

## Kodlar:

```
import nltk
from nltk.corpus import cmudict
from nltk.tokenize import word_tokenize, sent_tokenize

# cmudict ve punkt veri setlerini indir
nltk.download('cmudict')
nltk.download('punkt')

# cmudict'i yükle
d = cmudict.dict()

# Kelimenin hece sayısını hesapla
def count_syllables(word):
    try:
        # cmudict kullanarak kelimenin hece sayısını al
        return [len(list(y for y in x if y[-1].isdigit())) for x in
d[word.lower()][0]]
    except KeyError:
        # cmudict'te kelime bulunamazsa, hece sayısını tahmin etmek için ünlü
harfleri say
        return sum(1 for letter in word if letter in 'aeiouAEIOU')

# Flesch-Kincaid Okunabilirlik Seviyesi hesaplama
def flesch_kincaid_grade_level(text):
    # Metni cümlelere böl
    sentences = sent_tokenize(text)
    # Metni kelimelere böl
    words = word_tokenize(text)
    # Toplam hece sayısını hesapla
    syllable_count = sum(count_syllables(word) for word in words)
    # Cümle sayısını hesapla
    sentence_count = len(sentences)
    # Kelime sayısını hesapla
    word_count = len(words)

    # Flesch-Kincaid Okunabilirlik Seviyesi formülü
    grade_level = 0.39 * (word_count / sentence_count) + 11.8 * (syllable_count /
word_count) - 15.59
    return grade_level

# Okunabilirlik seviyesini belirleme
def determine_readability_level(grade_level):
    if grade_level < 6:
        return "İlkokul"
    elif 6 <= grade_level < 9:
        return "Ortaokul"
    elif 9 <= grade_level < 12:
        return "Lise"
    elif 12 <= grade_level < 16:
        return "Üniversite"
    else:
```

```
        return "Teknik Doküman"

# Analiz edilecek metin
text = """Burada örnek bir metin olacak. Metnin okunabilirlik seviyesini
belirleyeceğiz."""

# Okunabilirlik seviyesini belirle
grade_level = flesch_kincaid_grade_level(text)
readability_level = determine_readability_level(grade_level)
print("")
print("")
# Sonuçları yazdır
print(f"Metin: {text}")
print(f"Okunabilirlik Seviyesi: {readability_level} (Flesch-Kincaid Grade Level:
{grade_level:.2f})")
```