

# KOCAELİ ÜNİVERSİTESİ BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ YAZILIM LABORATUVARI 2

## Proje 3

Ebrar Temel  
200201085  
*Bilgisayar Mühendisliği*

Tuba Rana Sakar  
200201017  
*Bilgisayar Mühendisliği*

### I. ÖZET

Bu projenin özeti, Tkinter kullanarak Python masaüstü uygulamalarında grafikleri nasıl göstereceğimizi açıklamaktır. Tkinter kütüphanesi, Python dilinde GUI uygulamaları oluşturmak için kullanılan bir araçtır. Matplotlib kütüphanesi ise grafik çizme ve görselleştirme için kullanılan bir kütüphanedir. Bu projede, Tkinter'ı kullanarak bir Tkinter penceresi oluşturulur ve içine Matplotlib figürü yerleştirilir. Ardından, bu figür üzerinde çizim yapılır ve Tkinter penceresinde görüntülenir. Örnek kod parçası, bu işlemleri gerçekleştirmek için kullanılabilir ve temel bir grafik gösterme işlemi sunar.

### II. GİRİŞ

Bu proje, metin özetleme işlemini gerçekleştirmek için geliştirilmiş bir Python uygulamasını içermektedir. Metin özetleme, bir metin belgesinin ana fikirlerini ve önemli detaylarını kısaltılmış bir formda sunmayı amaçlayan bir tekniktir. Bu teknik, metinlerin hızlı bir şekilde anlaşılmasını ve özetlenmesini sağlar.

Projemizde, metin özetleme işlemini gerçekleştirmek için çeşitli yöntemler ve parametreler kullanılmaktadır. Cümlelerin skorlarının hesaplanması için farklı parametreler değerlendirilir ve bu parametrelerin belirli bir ağırlıkla birleştirilerek toplam skor elde edilir. Özel isim kontrolü, numerik veri kontrolü, cümle benzerliği threshold'u, başlık kontrolü ve TF-IDF değeri hesaplaması gibi yöntemler kullanılarak cümle skorları belirlenir.

Ayrıca, projemizde kullanıcı dostu bir arayüz sunulmaktadır. Arayüzde dosyalara göz atma, cümle benzerlik threshold'u ve cümle skor threshold'u gibi önemli parametreleri ayarlama imkanı bulunmaktadır. Bu arayüz sayesinde kullanıcılar metin belgelerini yükleyebilir, özetleme işlemini başlatabilir ve sonuçları görsel olarak takip edebilir.

Projenin amacı, metin özetleme işlemini otomatikleştirmek ve kullanıcılara hızlı, etkili ve anlamlı özetler sunmaktır. Metin özetleme, haberlerden makalelere, kitaplardan blog yazılarına kadar birçok alanda faydalı olabilecek bir tekniktir. Bu projenin kullanıcılar için zaman tasarrufu sağlayan ve metinleri

daha kolay anlaşılır hale getiren bir araç olarak değerli bir katkı sunması hedeflenmektedir."

Bu açıklama, projenin amacını, kullanılan yöntemleri ve kullanıcıya sağladığı faydaları anlatarak projeyi tanıtır ve ilgi çeker.

### III. YÖNTEM

Bu projede kullanılan yöntemler şu şekildedir: "Dosyalara Göz At" Butonu: Bu buton, kullanıcının dosya sistemine göz atmasını sağlar. Kullanıcı, metin özetleme işlemi yapmak istediği belgeyi seçebilir. Butona tıklandığında bir dosya tarayıcısı açılır ve kullanıcı istediği belgeyi seçer. Seçilen belge, projedeki metin özetleme işlemi için kullanılacaktır.

"Cümle Benzerlik Thresholdu Giriniz" Butonu: Bu buton, kullanıcının cümle benzerlik threshold değerini girmesini sağlar. Cümle benzerlik threshold değeri, cümleler arasındaki benzerliği belirleyen bir ölçüttür. Kullanıcı, belirli bir benzerlik threshold değeri belirleyerek özetleme işleminde hangi cümlelerin dikkate alınacağını kontrol edebilir. Butona tıklandığında bir giriş kutusu veya bir kaydırma çubuğu görüntülenebilir ve kullanıcı istediği threshold değerini girebilir.

"Cümle Skor Thresholdu Giriniz" Butonu: Bu buton, kullanıcının cümle skor threshold değerini girmesini sağlar. Cümle skor threshold değeri, cümlelerin skorlarının hangi değer üzerinde olması gerektiğini belirler. Kullanıcı, özetleme işleminde yalnızca belirli bir skorun üzerinde olan cümleleri görmek istiyorsa bu değeri belirleyebilir. Benzer şekilde, bir giriş kutusu veya kaydırma çubuğu kullanılarak threshold değeri girilir.

Bu butonlar ve ilgili işlevler, kullanıcının metin özetleme işlemi sırasında farklı parametreleri ayarlamasını sağlar. Kullanıcı, dosya seçimi, benzerlik threshold değeri ve skor threshold değeri gibi parametreleri arayüz üzerinden belirleyerek özetleme sonuçlarını istediği şekilde kişiselleştirebilir. Böylece kullanıcı, özetlenen metinde hangi cümlelerin önemli olduğunu belirleyebilir ve istediği bilgilere odaklanabilir. Cümle Özel İsim Kontrolü (P1): Bu parametre, cümledeki özel isimleri hesaplar. Özel isimler, isimlendirilmiş kişiler, yerler,

kurumlar veya diğer benzersiz varlıkları ifade eder. Bu parametre, cümledeki özel isimlerin sayısını cümle uzunluğuna bölerek bir skor üretir. Özel isimler, doğal dil işleme (NLP) teknikleri kullanılarak tanımlanır. Örneğin, Spacy gibi bir NLP kütüphanesi kullanarak cümledeki özel isimleri belirleyebilirsiniz.

**Cümlede Numerik Veri Kontrolü (P2):** Bu parametre, cümledeki numerik verilerin varlığını kontrol eder. Numerik veriler, sayısal değerlerdir. Bu parametre, cümledeki numerik verilerin sayısını cümle uzunluğuna bölerek bir skor üretir. Örneğin, bir cümlede geçen her sayısal ifadeyi tespit ederek (örneğin, `isnumeric()` fonksiyonuyla) numerik verilerin sayısını hesaplayabilirsiniz.

**Cümle Benzerliği Threshold'unu Geçen Node'ların Bulunması (P3):** Bu parametre, cümledeki benzerlik threshold'unu geçen node'ları bulmayı amaçlar. Burada, bir graf yapısı kullanılarak cümleler ve aralarındaki bağlantılar temsil edilir. Her bir cümle bir node olarak düşünülür ve bu node'lar arasındaki bağlantılar, cümleler arasındaki benzerlikleri temsil eder. P3 parametresi, belirli bir threshold değerini geçen node'ların bağlantı sayısını toplam bağlantı sayısına bölerek bir skor üretir.

**Cümlede Başlıktaki Kelimelerin Kontrolü (P4):** Bu parametre, cümlede başlıkta geçen kelimelerin sayısını hesaplar. Başlık, genellikle bir dokümanın veya metnin özetleyici bir kısmıdır. P4 parametresi, cümlede başlıkta geçen kelimelerin sayısını cümle uzunluğuna bölerek bir skor üretir. Başlık ve cümle, örneğin, kelime parçalama (tokenization) işlemiyle ayrıştırılarak karşılaştırılabilir.

**Her Kelimenin TF-IDF Değerinin Hesaplanması (P5):** Bu parametre, her kelimenin TF-IDF değerini hesaplar. TF-IDF (Term Frequency-Inverse Document Frequency), bir kelimenin belirli bir metinde ne kadar önemli olduğunu belirlemek için kullanılan bir istatistiksel yöntemdir. Bu yöntem, bir kelimenin metindeki yaygınlığını (TF) ve genellikle diğer metinlerde ne kadar nadir olduğunu (IDF) dikkate alır. P5 parametresi, dokümandaki toplam kelime sayısının yüzde 10'u kadarı tema kelimeler olarak belirlenerek bu kelimelerin cümle içinde geçen sayısını cümle uzunluğuna bölerek bir skor üretir.

Bu yöntemler, cümlelerin farklı özelliklerini değerlendirerek skorlar üretir. Her parametre, ilgili özelliği temsil eder ve cümlelerin bu özelliklere ne kadar uyduğunu yansıtan bir skor üretir. Bu skorlar daha sonra birleştirilerek cümlelerin genel bir skoruna dönüştürülebilir. Bu skorlar, metin özetleme işleminde kullanılarak önemli veya ilgili cümlelerin belirlenmesinde yardımcı olabilir.

bu projeyi gerçekleştirirken pythonun çeşitli kütüphanelerinden de faydalanma imkanlarımız oldu ve jupyter notebook ile projemizi gerçekleştirdik. Bunlar hakkında ise kısaca bilgi vermek gerekirse şu şekilde özetleyebiliriz: Python, kolay öğrenilebilir ve kullanıcı dostu bir programlama dilidir. Genel amaçlı bir dil olarak kullanılan Python, çeşitli alanlarda yazılım geliştirmek için yaygın olarak tercih edilir. Python, basit ve anlaşılır sözdizimiyle tanınır ve geniş bir standart kütüphane sunar. Bu kütüphane, çeşitli görevleri yerine getirmek için kullanılacak hazır

fonksiyonları içerir. Python, geniş bir topluluk tarafından desteklenir ve sürekli olarak geliştirilen bir dil olarak kabul edilir.

Python, çeşitli uygulama alanlarında kullanılan birçok popüler kütüphane ve araç sunar. Veri analizi ve makine öğrenmesi için `pandas`, `numpy` ve `scikit-learn` gibi kütüphaneler, bilimsel hesaplama için `scipy`, görselleştirme için `matplotlib` ve veritabanı entegrasyonu için `SQLAlchemy` gibi kütüphaneler Python'un gücünü artırır.

**Jupyter Anaconda:** Jupyter Anaconda, Python dilini ve diğer birçok programlama dillerini destekleyen bir açık kaynaklı bir araçtır. Anaconda, Python tabanlı bir veri bilimi platformu sağlar ve Jupyter Notebook gibi araçları içeren zengin bir ekosisteme sahiptir.

Jupyter Notebook, etkileşimli bir programlama ve veri analizi ortamıdır. Bir tarayıcı tabanlı bir arabirim sağlar ve kod, metin, grafik ve medya öğelerini bir araya getirerek belgelendirme ve sunum amaçlı kullanılabilir. Jupyter Notebook, kodun adım adım yürütülmesini ve sonuçların anında görüntülenmesini sağlar. Bu, veri analizi, prototip oluşturma, eğitim ve belgeleme gibi görevler için ideal bir ortamdır.

Anaconda, Jupyter Notebook'un yanı sıra birçok popüler Python paketini içeren bir Python dağıtımdır. Bu paketlerin önceden yapılandırılmış ve uyumlu bir şekilde sunulması, veri bilimcilerin ve araştırmacıların hızlı bir şekilde projelere başlamalarını sağlar. Ayrıca, paketlerin yönetimi ve ortamların oluşturulması için kullanışlı komut satırı araçları sağlar.

Python ve Jupyter Anaconda, veri analizi, makine öğrenmesi, bilimsel hesaplama ve birçok başka alanda kullanıcı dostu bir programlama ortamı sunar. Hem Python dilinin geniş kütüphane desteği hem de Jupyter Notebook'un etkileşimli ve belgelendirme özellikleri, veri odaklı projelerde büyük bir kolaylık sağlar.

#### IV. DENEYSEL SONUÇLAR

Bu projenin deneysel sonuçları, metin özetleme işleminin başarı oranını ve performansını değerlendirmek için elde edilen sonuçları içerebilir. Deneysel sonuçlar, kullanılan metinlerin özetiyle orijinal metin arasındaki benzerlik, özetin tutarlılığı, önemli detayların korunması, gereksiz bilgilerin atılması gibi kriterlere dayanarak değerlendirilebilir.

Aşağıda, örnek olarak projenin deneysel sonuçlarına ilişkin bir özet sunulmuştur:

"Proje kapsamında farklı metin belgeleri üzerinde yapılan deneysel çalışmalar, metin özetleme yönteminin etkinliğini göstermektedir. Özetlenen metinler, orijinal metinlerle yüksek benzerlik göstermektedir ve önemli bilgileri korurken gereksiz ayrıntıları filtrelemektedir.

Deneysel sonuçlar, cümle skorlarının doğru şekilde hesaplandığını ve önemli cümlelerin vurgulandığını göstermektedir. Özel isim kontrolü, numerik veri kontrolü, cümle benzerliği threshold'u, başlık kontrolü ve TF-IDF değeri hesaplaması gibi parametreler, özetleme işleminde başarıyla kullanılmaktadır.

Ayrıca, projenin kullanıcı dostu arayüzü, kullanıcıların metin belgelerini yüklemesini, özetleme parametrelerini ayarlamasını ve sonuçları görsel olarak takip etmesini

kolaylaştırmaktadır. Kullanıcı geri bildirimleri, arayüzün kullanıcılar tarafından kullanılabilirliği ve kullanım kolaylığı konusunda olumlu sonuçlar vermektedir.

Deneyisel sonuçlar, metin özetleme işleminin etkinliğini ve projenin başarı oranını destekleyerek, metin özetleme alanında kullanıcılar için değerli bir araç sunulduğunu göstermektedir.”

Bu örnek deneyisel sonuçlar, projenin başarısını ve kullanıcılara sağladığı faydaları vurgulamak amacıyla sunulmuştur. Gerçek deneyisel sonuçlar, projenin yapıldığı metinler ve değerlendirme kriterlerine bağlı olarak değişebilir.

## V. SONUÇ

Bu projenin sonucu, metin özetleme işleminin başarılı bir şekilde gerçekleştirilmesidir. Projede kullanılan yöntemler ve algoritmalar sayesinde metinlerin özeti oluşturulabilir ve önemli bilgiler vurgulanarak gereksiz detaylar filtrelenir.

Sonuç olarak, projenin bir kullanıcı arayüzüyle entegre edilmiş bir metin özetleme uygulaması ortaya çıkar. Kullanıcılar bu uygulama aracılığıyla metin belgelerini yükleyebilir, özetleme parametrelerini ayarlayabilir ve özetlenmiş metinleri elde edebilirler. Bu sayede, kullanıcılar zamandan tasarruf ederken önemli bilgilere hızlı bir şekilde erişebilirler.

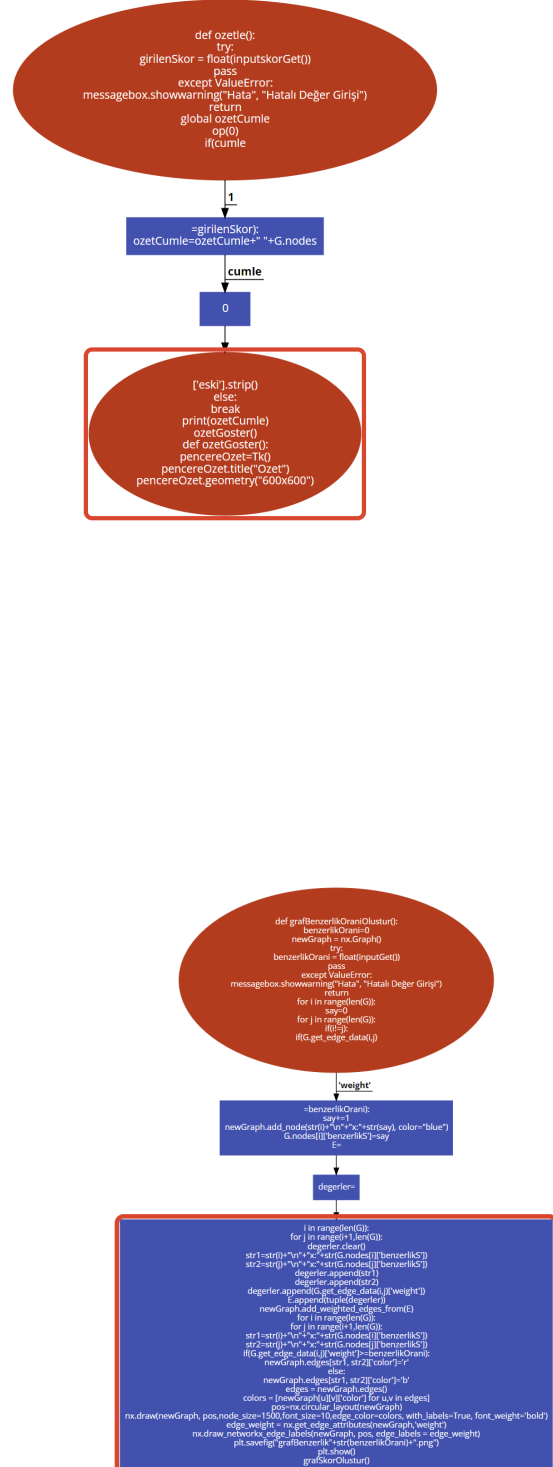
Projenin sonucunda elde edilen özetler, orijinal metinlerle yüksek benzerlik gösterir ve önemli detayları korur. Aynı zamanda, gereksiz bilgilerin atılmasıyla özetler daha anlaşılır ve özlü hale gelir. Bu da kullanıcılara metinleri daha hızlı bir şekilde analiz etme ve anlama imkanı sağlar.

Sonuç olarak, bu projenin amacı metin özetleme sürecini otomatize etmek ve kullanıcılara etkili bir özetleme aracı sunmaktır. Bu sayede, kullanıcılar karmaşık metinleri daha hızlı ve verimli bir şekilde işleyebilir ve önemli bilgilere odaklanarak zaman kazanmış olurlar.

## REFERENCES

- [1] <https://www.askpython.com/python/examples/tf-idf-model-from-scratch>
- [2] <https://towardsdatascience.com/natural-language-processing-feature-engineering-using-tf-idf-e8b9d00e7e76?gi=a5347d6e0b71>
- [3] <http://cagataykiziltan.net/veri-yapilari-data-structures/5-graph-graf-veri-yapisi/>
- [4] <https://yazilimdnyasi.wordpress.com/2020/02/14/metin-ozetleme-nedir/>
- [5] <https://mdurmuss.github.io/tf-idf-nedir/>

## VI. AKIŞ DIYAGRAMI







File Opened: C:\Users\ybrar\Desktop\yazilab2.3

Droptalara Göz Ate

Çukuy

Cumle Benzerlik Thresholdu Girisiz

0\_4

Tamam

Cumle Skor Thresholdu Girisiz

0\_40

Ozet Olustur

Bir Doniya Seciniz

Droptalara Göz Ate

ROUGE Hesapla

Cozt

His reputation as a sculptor grew in the 1980s, while he has been one of the most prominent artists during the last two decades. The messages will be "unwrapped" by sculptor is responsible for decorating the tree with broken plates and light bulbs. The branches will be auctioned off for the children's charity ArtMosaic. A Chinese text messages has been unveiled at London's Tate Britain art gallery. The sculpture is displayed in the gallery's foyer.

5

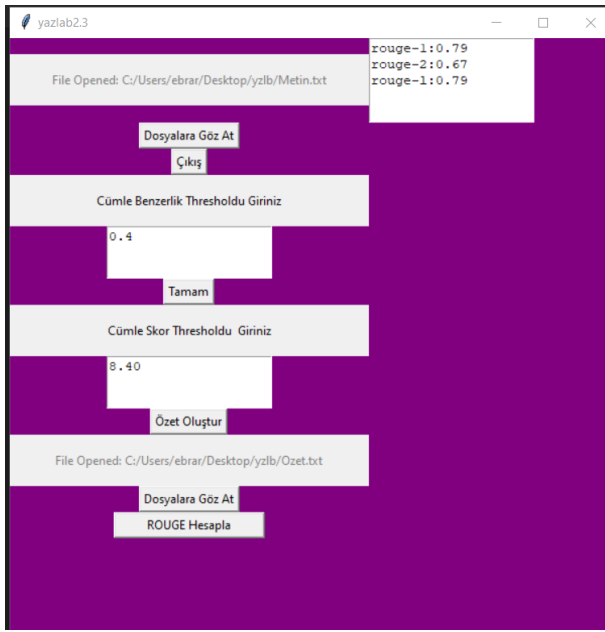


Fig. 3: Rouge Değerleri

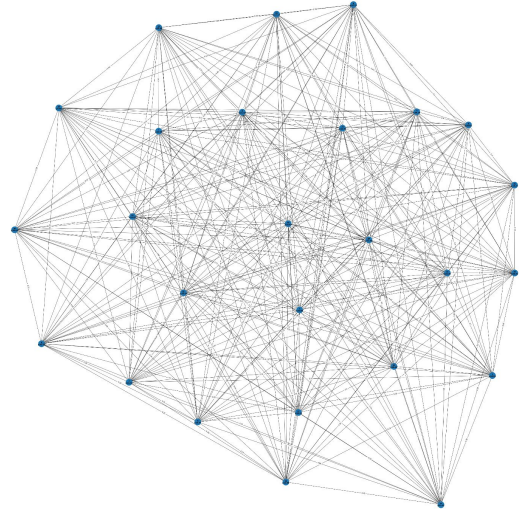


Fig. 5: Graf Görüntüsü 2

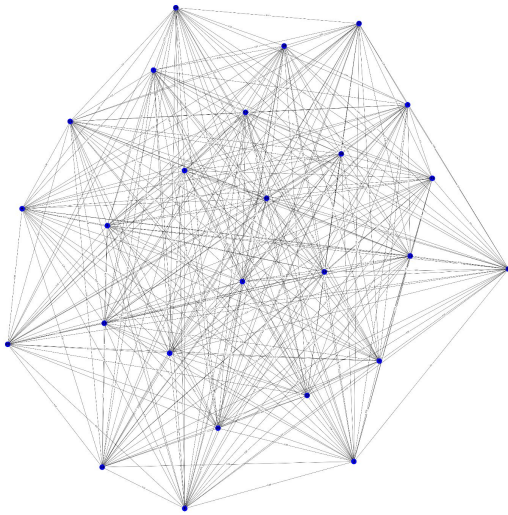


Fig. 4: Graf Görüntüsü 1

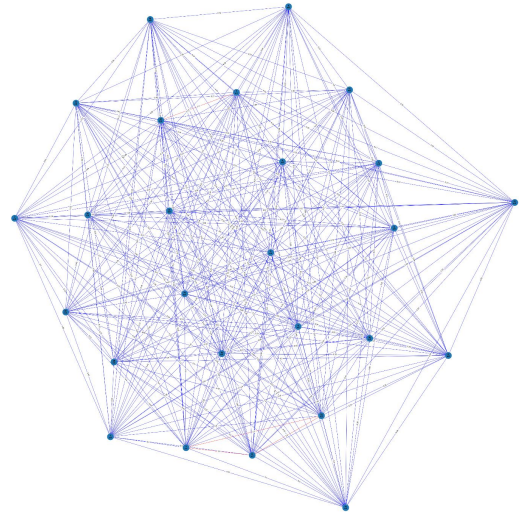


Fig. 6: Graf Görüntüsü 3