,

Bakırçay Üniversitesi

Bilgisayar Mühendisliği

Proje 3

3. Hafta

Hazırlayanlar

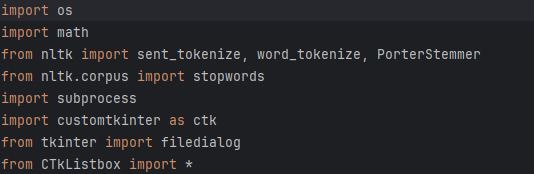
Talha Şükrü Üstün 230601091

Mehdi Dursun 230601025

Github Link: <https://github.com/Talhastn/Text_Summarization>

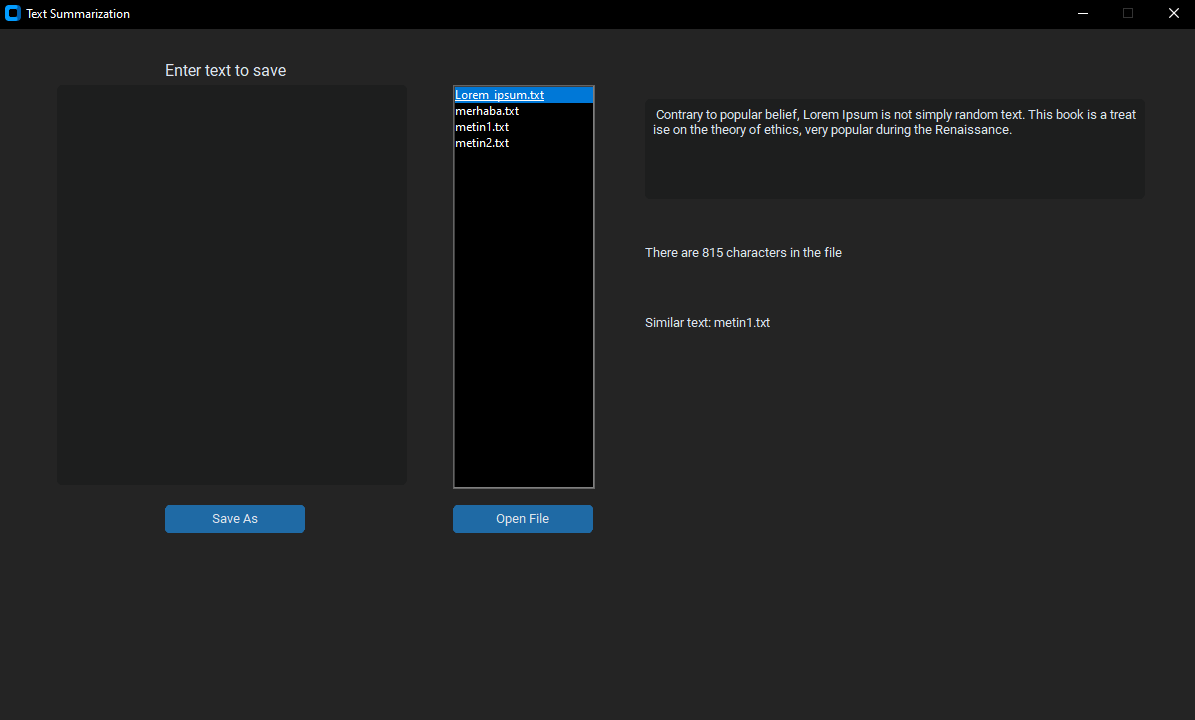
Hafta Genel Değerlendirmesi

Arayüz



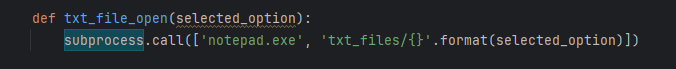
Projenin ikinci haftasında kullanılan kütüphaneler yukarıdaki gibidir.

Customtkinter, Tkinter ve CtkListbox kütüphaneleri kullanıcı arayüzü oluşturmak için seçilmiş ve kullanılmıştır. Arayüzün son hali aşağıdaki gibidir.



Kullanıcı istediği metni textbox’a yazabilecek ve bunu kaydedebilecek. Kaydedilen txt dosyası sağ tarafta bulunan listbox’a gidecek ve burada listelenecek. Her yeni dosya eklendiğinde listbox kendini yenileceyecektir. Eğer listbox içerisinde bulunan bir dosyanın adına tıklanır ve Open File butonuna basılırsa seçtiği dosya bilgisayarında seçili olan txt dosyalarını açma uygulamasında gösterilecek. Listbox’ın içerisinde yapılan her seçimin bilgileri listboxın hemen sağ tarafında gösterilecek. En üstte seçilen txt belgesinin özeti, altında, belgenin kaç adet karakterden oluştuğu, en altta ise listbox’ın içerisinde bulunan belgelerden seçilen belgeye anlam olarak en yakın dosyanın adı yazılmaktadır.

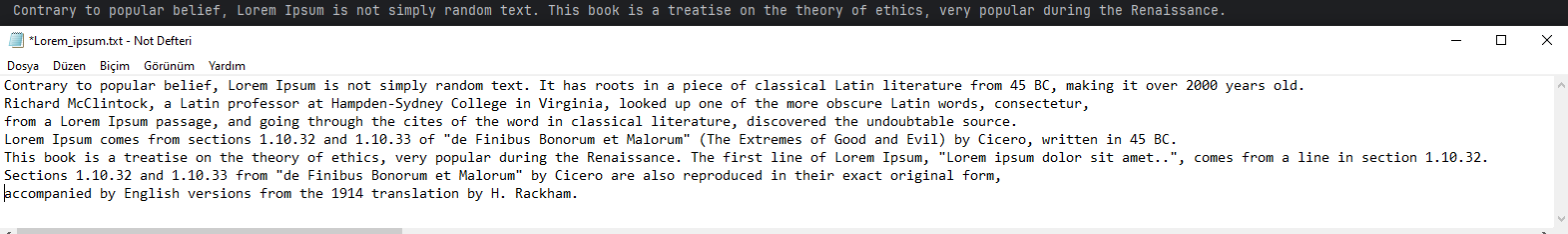
Kullanılan subprocess kütüphanesi listbox’ın içerisindeki txt dosyalarını açmak için kullanılmıştır.



Os kütüphanesi ise listbox’ın içerisine txt dosyalarını eklemek için kullanılmıştır.

Doğal Dil İşleme

Math ve nltk kütüphaneleri bir metnin özetini çıkarmak için kullanılmıştır.

Yukarıdaki resimde en yukarda siyah arka planda yazan yazı, altta duran txt dosyasının içerisindeki yazının kısa bir özetidir.

Program sadece ingilizce metinler üzerinde çalışmaktadır. Çünkü stopwordsler ingilizce olarak belirlenmiştir. Yazılan doğal dil işleme algoritmasında belirlediğimiz tokenlar a’dan z’ye kadar olduğu için metin içerisindeki hiçbir sayıyı özet içerisinde bulundurmamaktadır. Metnin özeti anlık olarak çıkarılmakta ve kullanıcıya arayüzde gösterilmektedir.