**BURSA TEKNİK ÜNİVERSİTESi**

**MÜHENDİSLİK VE DOĞA BİLİMLERİ FAKÜLTESİ**

**BİLGİSAYAR MÜHENDİSLİĞİ**

**ALGORİTMA VE PROGRAMLAMA DERSİ**

**PROJE ÖDEVİ**

**ALMANCA VEYA İNGİLİZCE DİLİNDE YAZILAN**

**METİNLERİN AYIRT EDİLMESİ**

**HAZIRLAYANLAR:**

**FURKAN MEMUS 19360859021**

**NESLİHAN ELİF ILGAZ 19360859061**

**ÖZET**

**8500 karakteri aşmayacak bir metin kullanıcı tarafından girilir.**

**Girilen metin yabancı karakterlerden temizlenir.**

**Bigram ve trigramlarin metinde kaç kez bulunduğu hesaplanır.**

**Verilen frekans değerleriyle hesaplama yapılır ve metnin dilinin Almanca mı yoksa İngilizce mi olduğu belirlenir**

**YAPILAN ÇALIŞMALAR**

**A)**

**Kullanıcıdan metin almak için gets fonksiyonunu kullandık**

**Metinde geçen bigram ve trigramların frekansını hesapladık**

**B)**

**Girilen metindeki yabancı karakterlerin yerine boşluk karakteri getiricek void filter\_str(char str[]) fonksiyonunu kullandık**

**Bigram ve trigramlar aranırken sorun çıkmaması amacıyla metni**

**Temizledikten sonra bütün harflerin küçük harfe dönüşmesini sağladık**

**C)**

**Her bir bigramın metinde kaç kez tekrar ettiğini bulmak için**

**void calculate\_frequencies\_bi(char str[]) fonksiyonunu kullandık**

**İç içe 3 for döngüsü ile bigramları taradık ve saydık.**

**Her bir bigramın tekrar sayısı matrise atanır.**

**Hesaplanan frekansaları calculated\_frequencies[l] dizisine sırasıyla**

**yerleştirdik.**

**D)**

**Her bir trigramın metinde kaç kez tekrar ettiğini bulmak için**

**void calculate\_frequencies\_tri(char str[]) fonksiyonunu kullandık.**

**Hesaplanan frekansaları 10. elemandan başlayacak şekilde**

**calculated\_frequencies[l-10] dizisine sırasıyla yerleştirdik.**

**E)**

**void calculate\_distances() fonksiyonu ile İngilizce ve Almanca dilleri için ayrı ayrı puanlama yaptık .**

**Distances[0] dizisine İngilizce için hesaplanan puanın kuvveti atanır.**

**Distances[1] dizisineyse Almanca için hesaplanan puanın kuvveti atanır.**

**Projede bizden öklid hesaplaması yapmamız tavsiye ediliyordu ama bu**

**hesaplama kısa metinlerde hatalı sonuç veriyordu.**

**Bu yüzden sorunu çözecek yeni bir yöntem geliştirdik.**

**F)**

**void detect\_lang() fonksiyonunda distances[0] ve distances[1]’in değerlerini**

**karşılaştırdık . İngilizcenin kuvvet puanını Almancanın kuvvet puanına**

**böldük ve sonuç 1’den büyük çıktığında metnin dilinin İngilizce olduğu**

**ekrana yazdırılıyor, değilse Almanca olduğu yazdırılıyor.**