**TF-IDF Yöntemiyle Özetten Anahtar Kelime Tahmini**

**TF= Term Frequency / Terim Sıklığı  
IDF= Inverse Document Frequency / Ters Döküman Sıklığı**

**Terim sıklığı=seçili kelimemizin, metin içinde bulunan toplam kelimeler sayısına bölümüdür.**

**Ters Doküman Sıklığı= Metinlerimizin kaçında kelimemiz var bunu gösterir. Toplam metin adetimizin kelimemizi içeren metin adetine bölümünün logaritmasıdır.**

**Aşamalarım:**

**İlk olarak dergipark.com dan dergileri çektim.Ardından bu dergilerden en çok makale bulunan “Akademik Bakış Uluslararası Hakemli Sosyal Bilimler Dergisi” adlı derginin verilerini (başlık,alt\_başlık, özet,anahtar\_kelime) .csv uzantısına çevirdim.Ardından pandas kütüphanesi ile bu verileri okudum ve bu projede bana lazım olacak kısımları yani özet ve anahtar kelimeleri bir sözlük tipinde ki değişkenime kaydettim.5 adımda sonuca ulaştım;**

**1)Metin Düzeltme ve StopWords Düzeltme Aşaması :İlk olarak 650 veri bulundurduğum kendi stopwords umu oluşturdum.Metindeki düzeltilmesi gereken yerleri re kütüphanesi ile düzeltip stopwords dosyam da bulunan kelimeleri metinden çıkardım.**

**2)Metnimi kelimeler bazında inceleyip, min ve max bulunma oranlarını da vererek bu standarda uyan kelimelerin kullanılma sayılarını buldum.**

**3)Metnimin tamamında geçen kelimeler arasında en çok kullanılan bi-gram ve uni-gram kelimeleri inceledim.**

**4**)**Scklearn kütüphanesini import ederek tf\_idf yöntemi ile anahtar kelime bulma işlemi için sıfırla bölmeyi yani sonsuzluğu engellemek için “smooth\_idf”=True yaptım.Ardından “IDF” işlemini etkinleştirmek için use\_idf=True yaptım.**

**from sklearn.feature\_extraction.text import TfidfTransformer**

**tfidf\_transformer=TfidfTransformer(smooth\_idf=True,use\_idf=True)**

**tfidf\_transformer.fit(word\_count\_vector)**

**5)Gerekli hesaplamalar yapıldıktan sonra,dergide ki 2. Verinin özeti için programımı çalıştırdım ve skoru en yüksek olan 5 anahtar kelimeyi output olarak aldım.**

**Gerçek Anahtar Kelimeler :** **Yükseköğretim ,öz yönetimli öğrenme**

**TF-IDF Yöntemi ile Bulduğumuz Anahtar Kelimeler:**

**öğrenmeye hazırbulunuşluk 0.33,**

**öz yönetimli öğrenmeye 0.33,**

**öz yönetimli 0.33,**

**yönetimli öğrenmeye hazırbulunuşluk 0.33,**

**yönetimli öğrenmeye 0.33**