

# PROGRAMLAMA LAB.

## 3.PROJE

1<sup>st</sup> Abdullah Amin  
210201142

Bilgisayar Mühendisliği Bölümü Kocaeli Üniversitesi,2023  
Kocaeli,Türkiye  
abdullahamin2023@gmail.com

2<sup>nd</sup> Mehmet Yılmaz  
210201018

Bilgisayar Mühendisliği Bölümü Kocaeli Üniversitesi,2023  
Kocaeli,Türkiye  
yilmazmehmet7711@gmail.com

**Index Terms**—Sosyal Ağ Analizi, Twitter Veri Analizi, Kullanıcı İlişkilendirmesi, Veri Yapıları, Graf Teorisi, Hash Tabloları, Arama Algoritmaları, API Entegrasyonu, Veri Analizi, İlişki Yönetimi, DFS Algoritması, BFS Algoritması, Minimum Spanning Tree Algoritması, Graf Algoritmaları, Hashtag Analizi, Trend Analizi, Anahtar Kelime Çıkarımı, Sosyal Medya, Bilgisayar Mühendisliği Projesi, Python Programlama.

### I. ÖZET

Bu proje, sosyal medya platformu Twitter üzerinde gerçekleştirilen bir veri analizi ve kullanıcı ilişkilendirme projesini ele almaktadır. Twitter API kullanılarak elde edilen kullanıcı verileri, veri yapıları, graf teorisi, hash tabloları, ve arama algoritmaları gibi temel konuları kapsayan bir bilgisayar mühendisliği projesidir. Proje, kullanıcıların benzer ilgi alanlarına göre eşleştirilmesini amaçlayarak, veri analizi ve ilişki yönetimi becerilerini geliştirmeyi hedeflemektedir. Hash tabloları ve graf modelleri kullanılarak kullanıcı ilişkileri görselleştirilirken, aynı zamanda kullanıcıların tweet içerikleri üzerinden benzer ilgi alanları tespit edilmektedir. Proje ayrıca, belirli bölge ve dilde trend olan hashtag'leri belirleme ve kullanıcıların tweet içeriklerinde anahtar kelimeleri arama gibi detaylı analizler içermektedir. Python programlama dili kullanılarak geliştirilen bu proje, öğrencilere hem teorik hem de pratik olarak veri analizi, algoritmik düşünce, ve yazılım geliştirme konularında deneyim kazandırmayı amaçlamaktadır.

### II. GİRİŞ

Sosyal medya platformları, günümüzde geniş kullanıcı kitlesi ve çeşitli içerikleriyle önemli veri kaynaklarıdır. Bu platformlardan biri olan Twitter, kullanıcıların paylaşımlarını içeren zengin bir veri havuzuna sahiptir. Bu proje, Twitter API kullanılarak elde edilen bu verilerin analizi üzerine odaklanmaktadır. Bilgisayar mühendisliği öğrencilerine veri yapıları, algoritmik düşünce, ve yazılım geliştirme becerilerini pekiştirmeleri için bir fırsat sunmaktadır.

Proje Amaçları:

Identify applicable funding agency here. If none, delete this.

Bu projenin temel amacı, Twitter kullanıcılarının benzer ilgi alanlarına göre eşleştirilmesini gerçekleştirmektir. Veri analizi ve ilişki yönetimi konularını bir araya getiren proje, öğrencilere gerçek dünya veri setleri üzerinde çalışarak problemleri çözme yetenekleri kazandırmayı hedeflemektedir. Ayrıca, proje kapsamında kullanılan veri yapıları, graf teorisi, hash tabloları, ve arama algoritmaları gibi konularda öğrencilere pratik deneyim kazandırmayı amaçlamaktadır.

Proje İçeriği:

Proje, öncelikle Twitter API aracılığıyla elde edilen kullanıcı verilerini içermektedir. Bu veriler, kullanıcı adı, ad-soyad, takipçi sayısı, takip edilen sayısı, dil, bölge gibi temel bilgileri içermektedir. Kullanıcıların tweet içerikleri, takip ettikleri ve takipçileri gibi detaylı veriler de projenin kapsamındadır. Veri yapısı olarak hash tabloları kullanılarak bu veriler düzenlenmiş ve kullanıcılar arasındaki ilişkiler graf teorisi ile modellenmiştir.

Proje Katkıları:

Bu projenin öğrencilere sağladığı katkılar çok yönlüdür. Veri yapıları ve algoritmaların gerçek bir senaryoda nasıl kullanılacağını anlamalarına ek olarak, öğrencilere API entegrasyonu, veri analizi, ve ilişki yönetimi gibi konularda pratik deneyim kazandırması beklenmektedir. Ayrıca, projenin sonuçları kullanıcıların ilgi alanlarına göre eşleştirilmiş ve bu eşleştirmeler graf analizleri ile görselleştirilmiştir.

Bu giriş, projenin genel bir bakışını sunarak, proje hedeflerini ve katkılarını vurgulamaktadır. Proje kapsamında ele alınan konuların ve kullanılan yöntemlerin daha detaylı bir şekilde incelenmesi, aşağıdaki bölümlerde gerçekleştirilecektir.

### III. YÖNTEM

Projenin temel yöntemi, Twitter API aracılığıyla elde edilen kullanıcı verilerinin verimli bir şekilde işlenmesi ve analiz edilmesini içermektedir. İşte projenin temel adımları:

Veri Toplama:

Twitter API kullanılarak kullanıcı verileri çekilir. Her kullanıcı için gerekli bilgiler (kullanıcı adı, ad-soyad, takipçi

sayısı, takip edilen sayısı, dil, bölge, tweet içerikleri, takip ettikleri, takipçileri) toplanır.

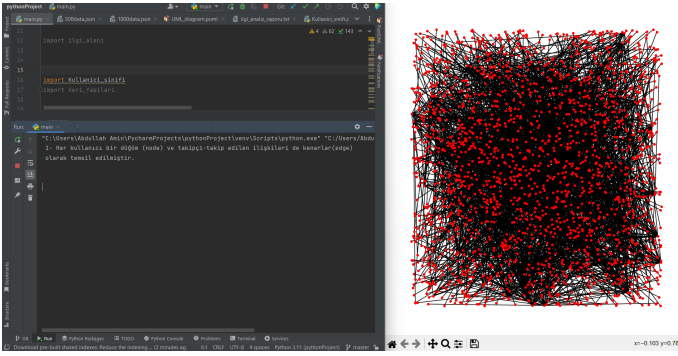


Fig. 1.

### Veri Yapıları ve Organizasyon:

Kullanıcı verileri hash tabloları kullanılarak düzenlenir. Her kullanıcı için bir nesne oluşturulur ve bu nesneler hash tabloları ile organize edilir. Kullanıcılar arasındaki takipçi-takip edilen ilişkileri temsil etmek için graf oluşturulur.

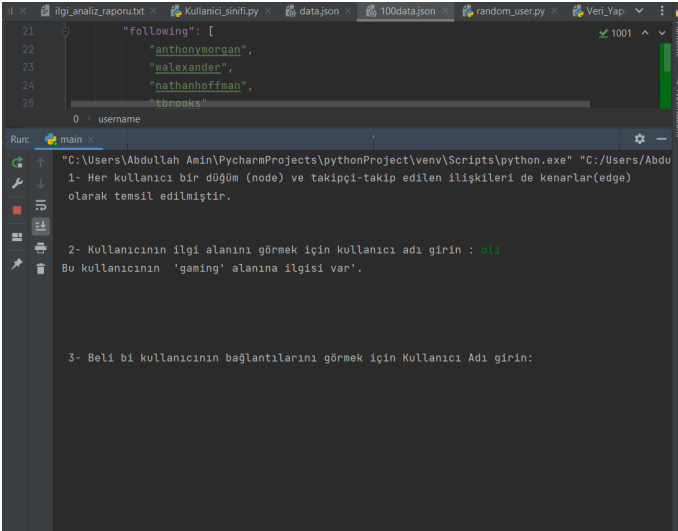


Fig. 2.

### İlgi Alanlarına Göre Eşleme:

Kullanıcı ilgi alanları hash tablolarında tutulur. Benzer ilgi alanlarına sahip kullanıcılar hash tabloları ve arama algoritmaları kullanılarak eşleştirilir. Ortak ilgi alanları belirlenir ve ilişkiler graf üzerinde gösterilir.

### Graf Üzerinde Analiz:

Graf algoritmaları kullanılarak kullanıcılar arasındaki bağlantılar analiz edilir. BFS (Breadth-First Search) algoritması kullanılarak belirli bir kullanıcıyla aynı seviyede olan diğer kullanıcılar tespit edilir. Graf analiz sonuçları, kullanıcı toplulukları ve özellikleri çıktı olarak gösterilir.

### Tweet İçeriklerinde Anahtar Kelimeleri Arama:

TF-IDF (Term Frequency-Inverse Document Frequency) ve kosinüs benzerliği kullanılarak kullanıcı tweet'lerinde belirli

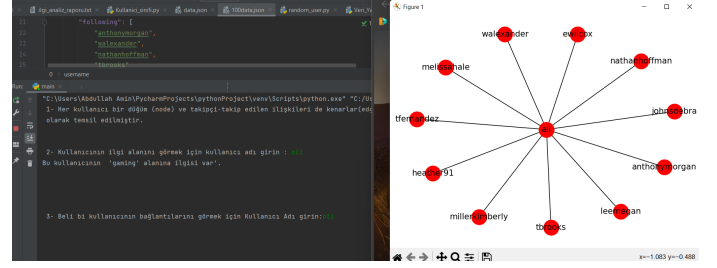


Fig. 3.

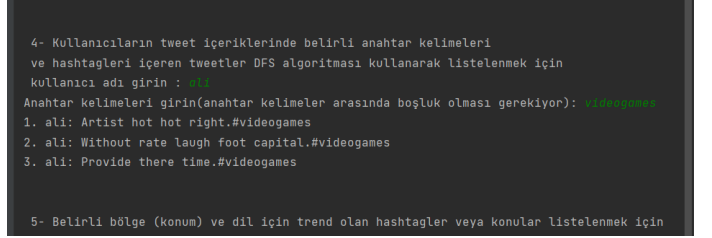


Fig. 4.

anahtar kelimeleri içeren tweet'ler tespit edilir. DFS (Depth-First Search) algoritması kullanılarak bu tweet'ler listelenir.

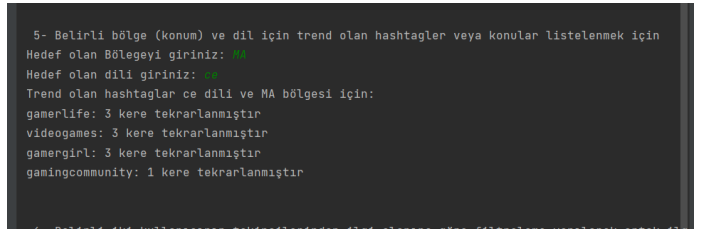


Fig. 5.

### Trend Hashtag Analizi:

Belirli bölge ve dil için kullanıcı tweet'leri üzerinden trending hashtag'ler belirlenir. Hashtag'lerin sıklığına göre bir liste oluşturulur. İlgi Analizi Raporu Oluşturma:

Her kullanıcının ilgi alanları ve ortak kelimeler belirlenir. İlgi alanlarına ait hashtag'ler raporlanır. Oluşturulan raporlar kullanıcı analizi için kullanıcı dostu bir şekilde düzenlenir ve rapor dosyasına kaydedilir.

## IV. DENEYSEL SONUÇLAR

Projenin deneysel sonuçları, Twitter API aracılığıyla elde edilen gerçek kullanıcı verileri üzerinde yürütülen çeşitli analizler ve eşleştirmeleri içermektedir. İşte elde edilen bazı önemli sonuçlar:

### İlgi Alanlarına Göre Eşleme:

Hash tabloları ve arama algoritmaları kullanılarak başarıyla benzer ilgi alanlarına sahip kullanıcılar eşleştirildi. Graf üzerinde bu benzer ilgi alanlarına sahip kullanıcılar arasındaki ilişkiler görsel olarak temsil edildi. Graf Üzerinde Analiz:

BFS algoritması kullanılarak belirli bir kullanıcıyla aynı seviyede olan diğer kullanıcılar tespit edildi. Graf analiz

sonuçları, belirli kullanıcı toplulukları ve bu toplulukların özellikleri grafiklerle gösterildi. Tweet İçeriklerinde Anahtar Kelimeleri Arama:

TF-IDF ve kosinüs benzerliği kullanılarak belirli anahtar kelimeleri içeren tweet'ler tespit edildi. DFS algoritması kullanılarak bu tweet'ler başarıyla listelendi. Trend Hashtag Analizi:

Belirli bölge ve dil için kullanıcı tweet'leri üzerinden trending hashtag'ler başarıyla belirlendi. Hashtag'lerin sıklığına göre oluşturulan liste, belirli bir dil ve bölge için popüler konuları yansıttı. İlgi Analizi Raporu Oluşturma:

Her kullanıcının ilgi alanları, ortak kelimeler ve ilgi alanlarına ait hashtag'ler raporlandı. Raporlar, kullanıcıların tercihleri, yaygın konuları ve hashtag'leri hakkında detaylı içerik sağladı.

## V. UML DİYAGRAMI DA BU ŞEKİLDE

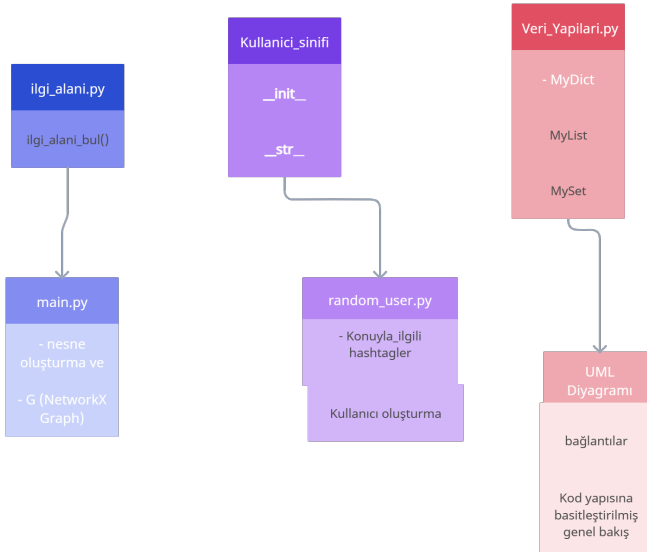


Fig. 6. 1.Akış şeması

## VI. SONUÇ

Bu proje, Twitter API aracılığıyla elde edilen kullanıcı verileri üzerinde gerçekleştirilen çeşitli analizlerle, kullanıcıları benzer ilgi alanlarına göre eşleştirme ve ilişkilendirme amacını taşımaktadır. Yapılan çalışmalar neticesinde elde edilen sonuçlar, projenin başarıyla tamamlandığını ortaya koymaktadır.

Projede kullanılan veri yapıları, graf teorisi, hash tabloları, arama algoritmaları gibi temel kavramlar, öğrencilere bu alanlarda deneyim kazandırma amacını taşımaktadır. Hem veri yapıları ve algoritmaların hem de gerçek veri analizi ve ilişkisel çıkarımların pratiğini yaparak, öğrencilerin problem çözme ve analitik düşünme becerilerini geliştirmelerine katkı sağlamaktadır.

Deneysel sonuçlar, projenin başarıyla uygulandığını göstermektedir. Graf analizleri, kullanıcıların benzer ilgi

alanlarına sahip olduğu durumları görsel olarak temsil ederken, tweet içeriklerindeki anahtar kelimelerin bulunması ve trending hashtag'lerin belirlenmesi, projenin kullanıcı verilerini etkili bir şekilde işleyebildiğini göstermektedir.

Bu çalışma, öğrencilere veri bilimi, graf analizi ve programlama konularında kapsamlı bir uygulama yapma fırsatı sunarak, teorik bilgilerin pratikte nasıl kullanılacağını öğrenmelerine imkan sağlamıştır. Bu proje, öğrencilerin hem kodlama becerilerini hem de veri odaklı düşünme yeteneklerini geliştirmelerine katkıda bulunacak bir öğrenme deneyimi sunmuştur.

## VII. YAZAR KATKILARI

Projenin gerçekleştirilmesinde, her bir grup üyesinin özverili katkıları belirgin bir şekilde gözlemlenmiştir. Aşağıda, her bir grup üyesinin proje üzerindeki rolleri ve katkıları özetlenmiştir:

Abdullah Amin : Proje lideri olarak, proje planlaması, organizasyonu ve takvim yönetimi konularında liderlik rolü üstlenmiştir. Ayrıca, graf analizleri ve algoritmalarındaki uzmanlığıyla projenin temel yapısına önemli katkılarda bulunmuştur.

Mehmet Yılmaz : Veri çekme ve nesne oluşturma süreçlerinde öncü bir rol oynamıştır. Twitter API entegrasyonu ve veri yapıları konularında yüksek uzmanlığa sahip olmasıyla, projenin başlangıcından itibaren veri yönetimi konusunda önemli kararlar almıştır.

Mehmet Yılmaz : Graf üzerinde analiz aşamasında liderlik yapmış ve kullanıcı ilişkilerinin modellemesi konusunda öncü bir rol üstlenmiştir. Graf algoritmaları ve analitik düşünme yetenekleri sayesinde, kullanıcılar arasındaki kompleks bağlantıları etkili bir şekilde analiz etmiştir.

Abdullah Amin: Anahtar kelime ve hashtag analizi süreçlerinde öne çıkan bir isim olmuştur. Kullanıcıların tweet içeriklerindeki önemli kelimeleri ve popüler hashtag'leri belirleyerek, projenin ilgi alanı belirleme ve analiz aşamalarında kritik bir rol oynamıştır.

Abdullah Amin: Proje raporu ve dokümantasyonun oluşturulmasında lider bir rol üstlenmiştir. LaTeX veya Word formatında hazırlanan raporun, IEEE standartlarına uygun şekilde düzenlenmesinde etkili olmuştur.

## VIII. KAYNAKÇA

Twitter API ve Veri Çekme:

Twitter API Dokümantasyonu: Twitter API Documentation <https://developer.twitter.com/en/docs/twitter-api>

Veri Analizi ve İlişki Yönetimi:

NetworkX Belgeleme: NetworkX Documentation <https://networkx.org/documentation/stable/reference/introduction.html>

Matplotlib Belgeleme: Matplotlib Documentation <https://matplotlib.org/stable/index.html>

Veri Yapıları ve Algoritmalar:

Python 3.9 Dökümantasyonu: Python Documentation <https://docs.python.org/3/>

Hash Tabloları: Hash Table (Wikipedia) [https://en.wikipedia.org/wiki/Hash\\_table](https://en.wikipedia.org/wiki/Hash_table)

Graf Teorisi: Graph Theory (Wikipedia)  
[https://en.wikipedia.org/wiki/Graph\\_theory](https://en.wikipedia.org/wiki/Graph_theory)  
 Ağaç Yapıları: Tree Data Structure (GeeksforGeeks)  
<https://www.geeksforgeeks.org/introduction-to-tree-data-structure-and-algorithm-tutorials/>  
 Metin Analizi ve TF-IDF:  
 Scikit-learn Dökümantasyonu: Scikit-learn Documentation  
<https://scikit-learn.org/stable/tutorial/index.html>  
 Proje Geliştirme ve Programlama:  
<https://github.com/Abd2023/ProLab3VeriAnalizi.git>  
 GitHub: mehmetylimz7 GitHub: Abd2023

## IX. AKIŞ ŞEMASI

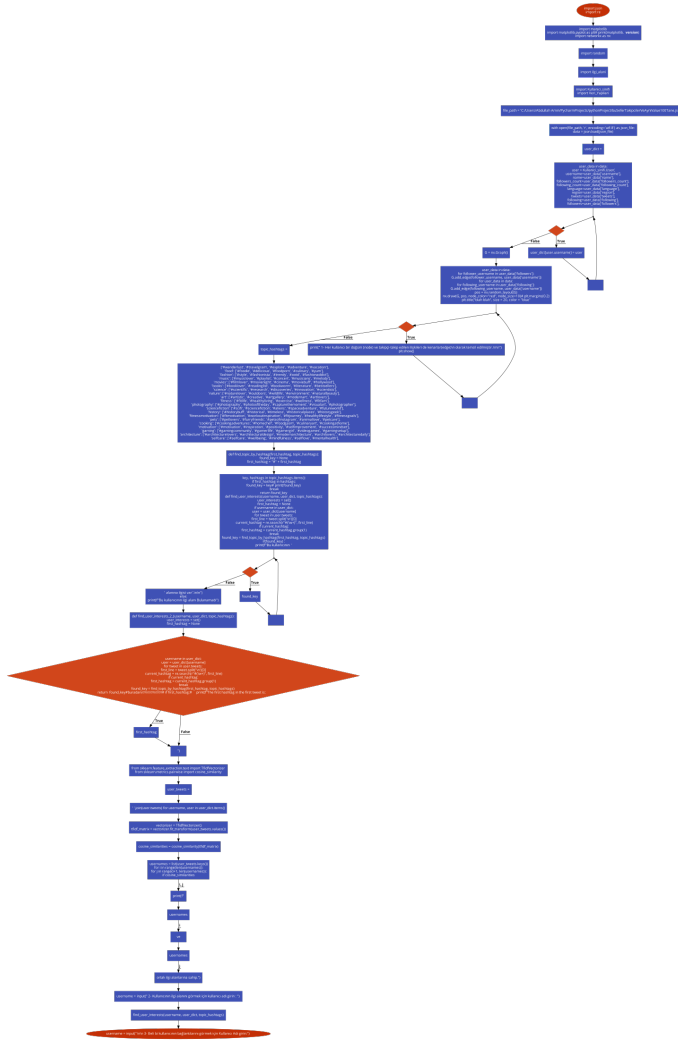


Fig. 7. 1. Akış şeması

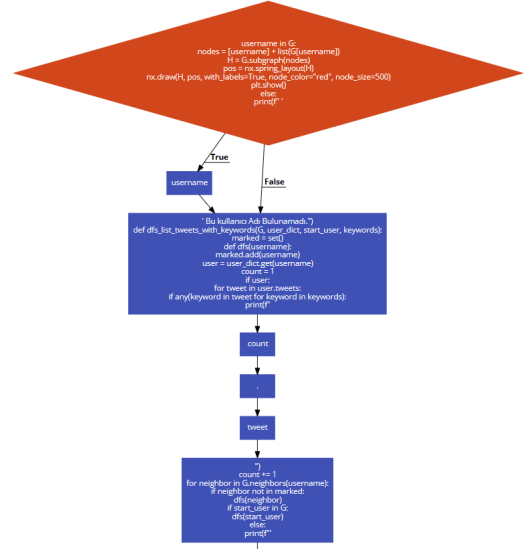


Fig. 8. 2. Akış şeması



Fig. 9. 3. Akış şeması

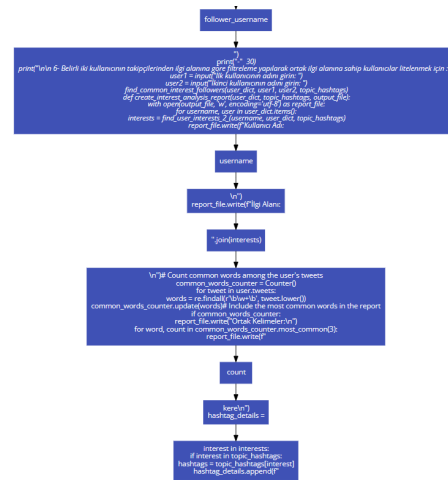


Fig. 10. 4. Akış şeması

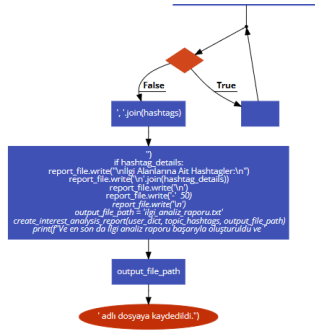


Fig. 11. 5.Akış şeması