

Kocaeli Üniversitesi

X (Twitter) Veri Analizi ve Kullanıcı İlişkilendirmesi

Berk Güven
Bilgisayar Mühendisliği
Kocaeli Üniversitesi
Kocaeli, Türkiye
berkconfidence@gmail.com

I. ÖZET

Bu proje, JSON aracılığıyla kullanıcı verilerini çekerek bu verileri analiz etmeyi ve kullanıcılar arasında benzer ilgi alanlarına göre eşleştirmeler yapmayı hedeflemektedir. Bununla birlikte kullanıcı verilerini, hash tablolarıyla organize ederek ve arama algoritmalarını kullanarak verimli bir şekilde ilgi alanlarına göre eşleştirme yapmayı amaçlamaktadır.

II. ÇÖZÜM

Projeye öncelikle verileri çekme aşaması ile başlanmaktadır. Bu kısımda veriler JSON dosyasından çekiliyor. Bu dosyada on bin kullanıcı, tweetler, takipçi-takip edilen kullanıcılar, bölge, dil vb şeyler bulunuyor. Javanın kütüphaneleri sayesinde dosyayı okuyarak bu veriler çekiliyor. Bu verileri ilk başta doğru şekilde çekebilmek için list yapısı kullanıyoruz. Daha sonra verileri(kullanıcılar) list yapısından BağlıListe classından tanımladığımız linkedlist nesnesi ve Graph classından tanımladığımız nesneye aktarıyoruz. Daha sonra bağlı listeye aldığımız bu kullanıcıları, İlgiAlanıKullanıcı classından tanımladığımız nesneyi kullanarak kullanıcıları hashtablosu sayesinde, tanımladığımız anahtar kelimeler ve ilgi alanları aracılığıyla tabloya kaydediyoruz.

Projedeki arayüz kısımdan önce son kısım olan analiz raporunda ise BağlıListe ve İlgiAlanıKullanıcı classlarının nesnelerini kullanıyoruz. Bu kısımda dosya oluşturmak için gerekli kütüphane ve değişkenleri tanımlıyoruz. Raporun nasıl olacağını hakkında bir taslak hazırladıktan sonra kodlama kısmına geçiyoruz. Öncelikle analiz ettiğimiz ilgi alanlarını yazdırıyoruz. Daha sonra bu ilgi alanlarında kaç kullanıcı olduğu ve anahtar kelimeleri yazdırıyoruz. Son olarak ise sırasıyla kullanıcıların ismini, ilgi alanlarını ve o ilgi alanlarına ait tweetlerini yazdırıyoruz.

Son kısımda ise işlemleri yapmak ve göstermek için arayüz oluşturuyoruz. Bu arayüz 5 kısımdan oluşuyor. İlk kısımdan başlarsak bizi “İlgi Alanı Kullanıcı” kısmı karşılıyor. Burada

öncelikle bir kullanıcı ve ilgi alanını seçiyoruz. Buna bağlı olarak bu ilgi alanına sahip diğer kullanıcıları listeliyoruz ve Bu kullanıcıların tweetlerini gösteriyoruz. İkinci kısım olan “Ortak İlgi Alanı Bulma” kısmında 2 adet kullanıcı seçimi yapmamız gerekiyor. Kullanıcıları seçtikten sonra bu iki kullanıcının ortak ilgi alanları listeleniyor ve seçilen ilgi alanına göre iki kullanıcının da tweetleri gösteriliyor. Üçüncü “Takipçileri Eşleştirme” kısmında 2 kullanıcı ve ilgi alanı seçiyoruz daha sonra ara butonu ile bu iki kullanıcının takipçilerinden ilgi alanına sahip kullanıcılar ve tweetleri listeleniyor. Dördüncü “Karşılaştırma” kısmında bizi 5 seçenek karşılıyor, bunlar: takipçi sayısı, takip sayısı, kullanılan dil, bölge ve tweet sayısı. Burada istenilen kullanıcı seçildikten sonra bu seçeneklerle aynı değere sahip diğer kullanıcılar ve bilgisi listeleniyor. Son kısımda ise “Analiz Raporu” adlı butona basarak ise projenin bulunduğu dosyaya txt dosyasını oluşturuyoruz.

III. YÖNTEM

Bu kısımda, çözüm aşamasında projenin genel hatlarla açıklanmış hali teknik kısımlarıyla açıklanacaktır. Proje boyunca kullanılacak olan 4 nesneyi main metodunda tanımlıyoruz.

- ObjectMapper mapper = new ObjectMapper();
- BağlıListe linkedList = new BağlıListe();
- Graph graph = new Graph();
- İlgiAlanıKullanıcı alan = new İlgiAlanıKullanıcı();

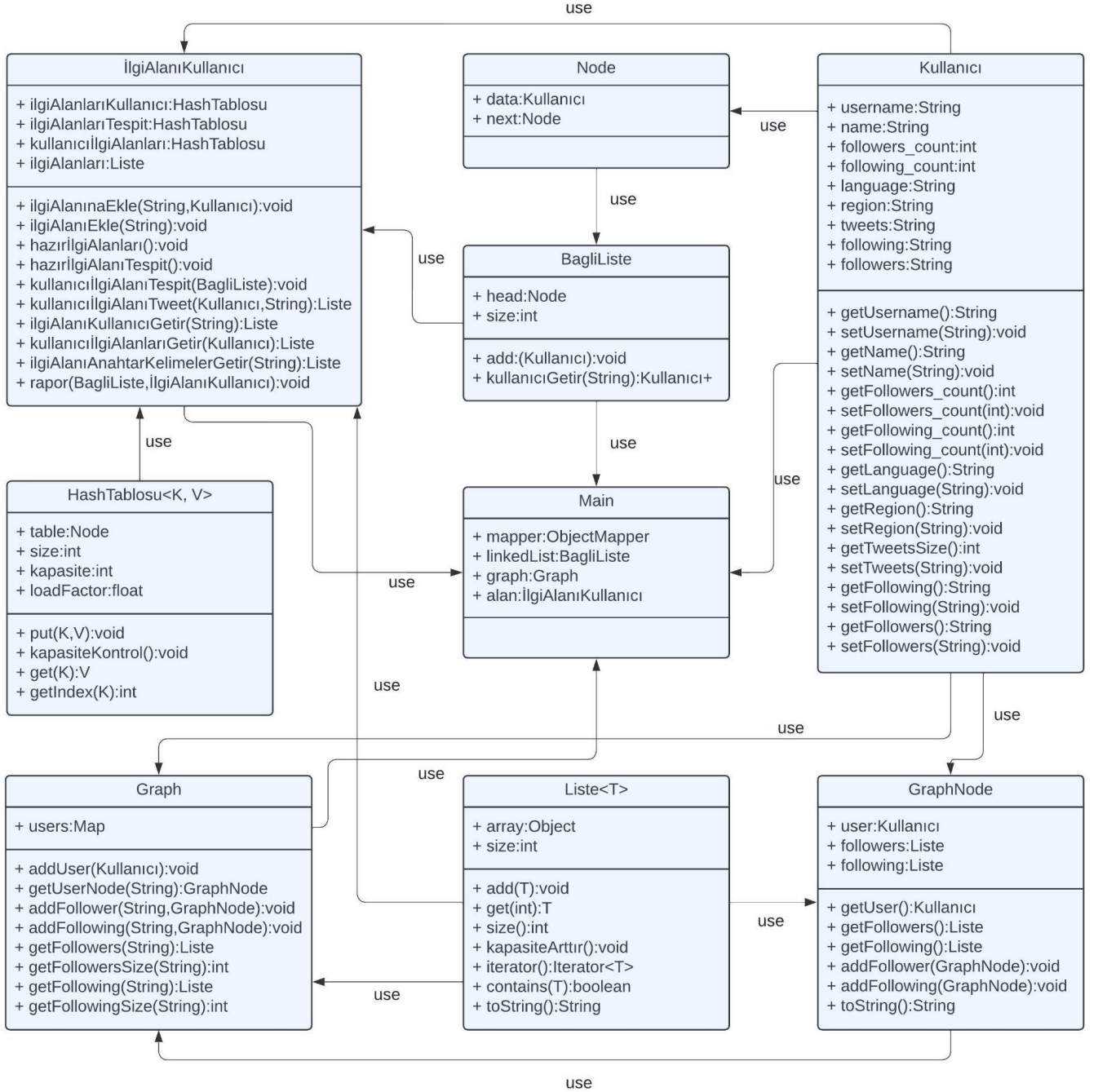
alan nesnesi ile “hazırİlgiAlanları” ve “hazırİlgiAlanıTespit” metodlarını İlgiAlanıKullanıcı classından kullanıyoruz. Bu metodlar hazır ilgi alanlarını ve anahtar kelimelerini hashtabllarına ekliyor.

- List<Kullanıcı> kullanıcıList ile verileri çekiyoruz.

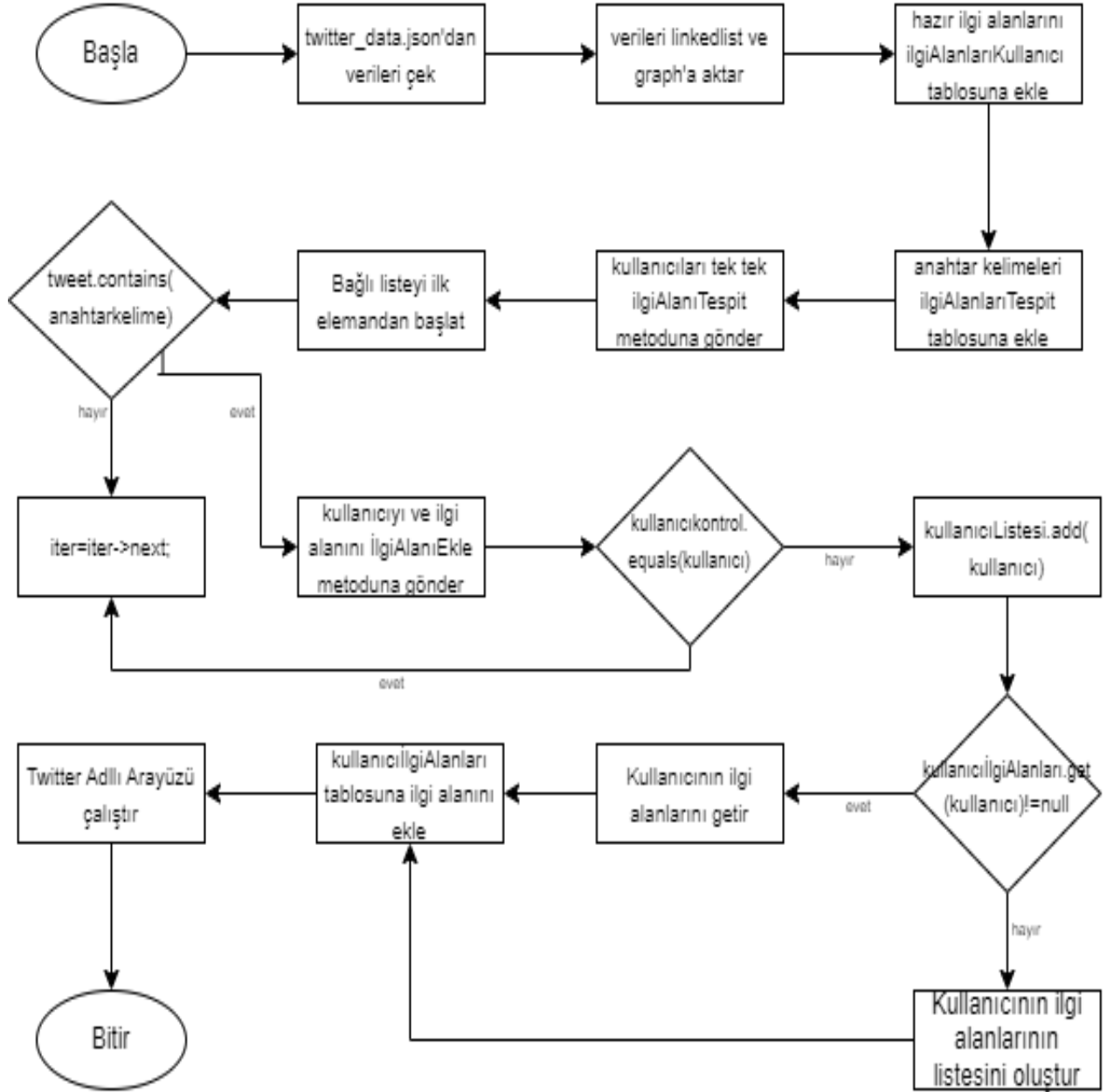
Bu verileri ise linkedList’e ve graph’a aktarıyoruz. Daha sonra linkedList’e aktardığımız verileri alan nesnesi ile

kullanıcıİlgiAlanıTespit metodu kullanarak her kullanıcının tweetlerini inceleyerek ona uygun ilgi alanını bulup hashtablosuna ekliyoruz. Daha sonra arayüzü oluşturup çözüm kısmında anlatılan 5 kısmı, uml diyagramında gösterilen metotlar sayesinde çözüyoruz.

IV. UML DİYAGRAMI



V. Akış Şeması



VI. Arayüz Çıktıları

Twitter İşlem Paneli

- İlgi Alanı Kullanıcı
- Ortak İlgi Alanı Bulma
- Takipçileri Eşleştirme
- Karşılaştırma
- Analiz Raporu

İlgi Alanına Göre Kullanıcı Bulma

Eşleştirmek İstedğiniz Kullanıcı

Emirhan Örgü

İlgi Alanı Seçimi

Spor

Göster

Eşleşen Kullanıcı

Ege Gönütaş

Ara Eşleştir

Birinci Kullanıcının Tweetleri

- Aynı yıl İspanya Kraliyet Futbol Federasyonu tarafından ülkenin en iyi hakemi olarak gösterildi.

İkinci Kullanıcının Tweetleri

- Cibuti Futbol Federasyonu tarafından yönetilir.
- Venezuela Futbol Federasyonu tarafından düzenlenmektedir.

Twitter İşlem Paneli

- İlgi Alanı Kullanıcı
- Ortak İlgi Alanı Bulma
- Takipçileri Eşleştirme
- Karşılaştırma
- Analiz Raporu

Ortak İlgi Alanı Bulma

Birinci Kullanıcıyı Seçiniz

Meriç Akyüz

İkinci Kullanıcıyı Seçiniz

Ece Özgörkey

Ortak İlgi Alanı

Müzik

Ara Eşleştir

Birinci Kullanıcının Tweetleri

- Şarkıcının Aşk Melekleri albümünden sonra yayınlandığı albümdür.

İkinci Kullanıcının Tweetleri

- Albüm bir başan elde edemedi.
- Albümün kapağında kırmızı ve siyah renklerle "Evanescence" logosu var.

Twitter İşlem Paneli

- İlgi Alanı Kullanıcı
- Ortak İlgi Alanı Bulma
- Takipçileri Eşleştirme
- Karşılaştırma
- Analiz Raporu

Takipçileri Eşleştirme

Birinci Kullanıcıyı Seçiniz

Rüya Tütüncü

Birinci Kullanıcının Takipçileri

Emirhan Örgü

İkinci Kullanıcıyı Seçiniz

Berke Ekici

İkinci Kullanıcının Takipçileri

Ece Polat

Eşleştir

Ortak İlgi Alanı

Bilim

Ara

Birinci Kullanıcının Tweetleri

- Kimya Mühendisliği Bölümünü bitirdi.

İkinci Kullanıcının Tweetleri

- Yalnız okuyucunun az da olsa fizik eğitimi almış olması gerekir.
- Bunu biraz matematiksel terimlerle açıklamak da mümkündür.
- Kimyasal özellikleri tungsten ile benzerlik gösterir.

Twitter İşlem Paneli

- İlgi Alanı Kullanıcı
- Ortak İlgi Alanı Bulma
- Takipçileri Eşleştirme
- Karşılaştırma
- Analiz Raporu

Kullanıcı Karşılaştırma

Kullanıcı Seçiniz

Berk Aydın

Kullanıcı Seçiniz

Polat Başoğlu

Kullanıcı Seçiniz

Ege Elçiboğa

Takipçi Sayısı Eşit Olanlar

Yağız Akal

Takip Sayısı Eşit Olanlar

Dr. Rüya Çörekçi

Aynı Dile Sahip Kullanıcılar

Alp Bademci

Takipçi Sayısı

250

Takip Edilen Sayısı

303

Kullanılan Dil

tk

Ara

Ara

Ara

Kullanıcı Seçiniz

Dr. Ege Taşçı

Kullanıcı Seçiniz

Ada Tokgöz

Aynı Bölgeden Kullanıcılar

Ali Yıldızoğlu

Tweet Sayısı Eşit Kullanıcılar

Ebru Korol

Bölge

BQ

Tweet Sayısı

62

Ara

Ara