|  |  |
| --- | --- |
| **Proje adı** | **Twitter üzerinden elde edilen veriler kullanılarak, tweet duygu analizinin yapılması** |
| **Proje Grubu** | Öğrenci No: 16008118040  Öğrenci İsim Soyisim: Engin Karataş  Öğrenci Email: 16008118040@ogr.bozok.edu.tr |
| **Proje Özeti** | Bir konu hakkında twitter üzerinde insanların genel kanısı bilimsel yaklaşımlarla araştırmak istenildiğinde, ilgili konunun hastag`indeki ilgili tweet`leri teker teker okumak ve araştırma yapan araştırmacının bireysel fikirleri ile değerlendirmek gerekmektedir. Bu yaklaşım tamamen nesnel bir bakış açısına sahip olamayacağından, insanların kanılarının objektif bir şekilde değerlendirilmesini zorlaştıracaktır. Elastic search kullanarak Twitter tweet analizi projesiyle kullanıcıların yaptığı yorumlar bilgisayar tarafından analiz edilip konu bazında insanların genel kanılarının objektif bir şekilde değerlendirilmesi amaçlanmıştır.  Twitter`ın açık kaynak apisini kullanarak, tweet analizi gerçekleştirilecektir. Girdi olarak twittler twitter platformunda yazdıkları en fazla 280 karakter uzunluğundaki metin veriler alınacak ve çıkış olarak insanların konu/konular hakkındaki düşünceleri analiz edilecektir.  Sonuç kısmında ilişkisel olmayan elastic search veritabanı üzerinde kayıtlı analiz sonucu verilerinin, kibana üzerinde görselleştirilmesi hedeflenmektedir.  Analiz edilen verilerden elde edilen anlamlı veriler, Twitter api üzerinden twitter üzernde paylaşılmak üzere Twitter api`ye başvuru yapılan profil hesabından public account (herkese açık hesap) üzernden paylaşılacaktır. |
| **Çözüm Alternatifleri** | Konuyla ilgili akademik ve akademik olmayan şekilde farklı çalışmalar vardır. Yapılacak olan çalışmada full text search özelliği bulunan web 3.0`ın en iyi arama-veritabanı teknolojilerinden biri olan elastic search ve kibana kullanılacak olup, ilgili veriler, son kullanıcı için arama ve filtreleme kolaylığı sağlayacaktır.   1. Eyüp Sercan AKGÜL, Caner ERTANO, Banu DİRİ - Twitter verileri ile duygu analizi - <https://www.acarindex.com/pdfs/56610> 2. HARRAN ÜNİVERSİTESİ MÜHENDİSLİK DERGİSİ – Twitter`da duygu analizi <https://dergipark.org.tr/tr/pub/humder/issue/56545/772929> 3. Elasticsearch İle Twitter Analizi Bölüm-3 Duygusal Analiz - <https://www.cozumpark.com/elasticsearch-ile-twitter-analizi-bolum-3-duygusal-analiz/> |
| **Kullanılması düşünülen programlama dilleri ve üçüncü parti araçlar** | Programlama dili Python  **Neden python?**  Python metin madenciliği, makine öğrenmesi hesaplamalarındada oldukça kullanılan bir programlama dilidir. Anlaşılır ve popüler bir dil olan pythonun içerisinde, proje ihtiyacına yönelik olan kütüphanelerden yararlanılması kararlaştırılmıştır.  Bulut Tercihi AWS  **Neden AWS?**  Bulut tabanlı, Web 2.0 üzerinde çalışan, istenilen miktardaki donanımları kiralayabildiğimiz, bir çok teknolojik yeniliği içerisinde bulunduran bir platform olduğu için aws tercih edilmiştir.  Veritabanı Elasticsearch  **Neden Elastic Search?**  Elastic search dağıtık sistemlerde çalışabilen ilişkisel olmayan unstructered bir veri tabanıdır. Projede elastic search kullanılacaktır. Elastic search`in görselleştirme araçlarından faydalanılabileceği öngörülmüştür. Elasitc search indexlemeleri çok kuvvetli olduğundan performanslı sorgular yazılacaktır.  Önyüz teknolojileri Kibana ve Angular  **Neden kibana?**  Elastic search ile uyumlu çalışmaktadır.  **Neden Angular?**  Angular ile kullanıcı deneyimini ve veri sayısını zenginleştirmek için farklı şema ve grafikler oluşturabilmektedir.  Servis Twitter Api  **Neden Twitter api?**  Twitter api, big data olarak nitelendirilebilecek twitter`daki verilere erişilmesini sağlayacaktır. Bu sayede ülkemizde de oldukça kullanılan bir platform içindeki tweet`lere erişilecektir. |
| **Aylık çalışma planları** | |  |  | | --- | --- | | **Ay** | **Amaçlar** | | Ekim | Projeyedeki kullanılacak araçların kararlaştırılması ve kullanılacak teknolojilerin sistematik şekilde tasarlanmasının ardından ilk çalışmalara ve testlere başlanması. | | Kasım | AWS (Amazon Web Service) teknolojilerinin araştırılması. Pythonda, duygu analizi için kullanılacak yöntemlerin ve teknolojilerin araştırılması. Twitter api başvuru süreçlerinin başlatılması. | | Aralık | AWS sanal makine kiralanması için gerekli çalışmaların yapılması. Python3 programlama dilinin AWS`nin hizmeti içerisindeki OS`a kurulması ve secure SSH bağlantısı ile bağlanılması işlemi. Bağlanılan işletim sisteminde gerekli konfigrasyon ayarları gerçekleştirildikten sonra python üzerinde derleme ve hedef kod üzerinde çalışmaların yapılması. | | Ocak | Pyton dilinin yazım kurallarına alışkanlık kazanabilmek için python dilinin öğrenilmesi. Python üzerinde gerçekleştirilmesi planlanan algoritmaya karar verdikten sonra, kullanılacak yöntemler, methodlar ile ilgili kütüphane ve algoritmalarda geliştirme yapabilecek düzeyde çalışmalar yapılması. Twitter api bağlantısının gerçekleştirilmesi adına, Python`dan Twitter api`ye bağlantı sağlanması ve proje klasörünün oluşturulması. | | Şubat | Elasticsearch database üzerinde temel kavramların öğrenilmesi ve twitter`dan getirilen verilerin elastic database üzerine nasıl yazılacağının araştırılması. Elastic search üzerinde yazılan twitter verilerinin kontrol edilmesi ve veri tipleri ilgili çalışmalar yapılması. Ardından index, type, document, fields ve mapping gibi anahtar kelimelerin öğrenilmesi. | | Mart | Kibana üzerinde, elastic search ile alınmış verilerin görselleştirilmesinin temel düzeyde öğrenilmeye başlandıktan sonra kibana üzerindeki uygulanan arama algoritmasının nasıl çalıştığının araştırılması. Kibana üzernde gösterilecek verilerin hangi tür grafikte nasıl gözükeceğinin kararlaştırılması | | Nisan | Python üzerinde analiz edilen ve elastic search kibana üzerinde görselleştirilen anlamlı sonuçları, twitter api kulanılarak twitter aracığıyla paylaşma, yayınlama çalışmalarına başlanması. | | Mayıs | Projede kullanılan kodların refactoring işlemlerinin yapılması, uygulamanın ihtiyaç halinde frontend teknolojileri kullanılarak grafik arayüzü geliştirilemsi. | |
| **Riskler** | Çözülmek istenilen problemlerin ve projenin geliştirme aşamaları sırasında promlem meydana getirmesi olasıdır. Projenin çalışmasını engelleyecek şekilde büyük problemlerin çıkması durumunda meydana gelebilecek olan büyük problemler tahmin edilmiş ve ek geliştirme adımları oluşturulmuştur ve altarnatif çözümlere başvurulacaktır.  Twitter api üzernden twitter hesabında post işlemiyle projede sonuç olarak çıkarılan anlamlı istatistiklerin paylaşılamadığı durumda, bir web server üzernde veya bir text dosyası üzernden bu bilgiler yayınlanacaktır. Kibana üzerinde görselleştirilmesi planlanan değerler, kriterler dışında kalırsa, Angular gibi diğer frontend teknolojileri kullanılarak bu sorun giderilmesi planlanmıştır. |