



**Bilgisayar Mühendisliği Bölümü**

**Bitirme Projesi**

**Proje Başlığı**

**Öğrenci Ad-Soyad**

**Öğrenci Numarası**

**Danışman**

### Proje Bilgisi

Projenin Adı	
Toplam Sayfa Sayısı	

Sürüm	Yazarlar	Yorum	Tarih
İlk taslak			
Son Taslak			
Versiyon 1			
Versiyon 2			

### Yayın Onayı

\_\_\_\_\_  
**Ad Soyad**  
Tez danışmanı

\_\_\_\_\_  
**Ad Soyad**  
Tez koordinatör

\_\_\_\_\_  
**Ad Soyad**  
Tez Koordinatör  
Yardımcısı

# İÇİNDEKİLER

<b>1. Giriş .....</b>	<b>1</b>
1.1. Projenin amacı.....	1
1.2. Projenin Kapsamı, Çıktıları ve Tablolar.....	1
1.3. Projenin Hedefleri ve Başarı Kriterleri .....	3
1.4. Proje Üyelerinin Görev ve Sorumlulukları.....	3
<b>2. Yapılan Çalışma .....</b>	<b>4</b>
2.1. Makine Öğrenmesi .....	6
2.2. Sınır Ağı .....	6
2.3. Bayes Algortiması .....	6
2.4. Sözlük Tabanlı Sistem .....	6
<b>3. Önerilen Sistem .....</b>	<b>7</b>
3.1. Genel Bakış .....	7
3.2. Analiz ve Görselleştirme Araçları .....	8
3.3. Önerilen Sistem Mimarisi .....	8
3.4. Fonksiyonel Gereksinimler .....	8
3.5. Fonksiyonel Olmayan Gereksinimler .....	9
3.5.1. Kalite Gereksinimleri .....	9
3.5.2. Performans Gereksinimleri .....	10
3.6. Gerçek Olmayan Gereksinimler .....	10
3.7. Sistemdeki Modeller .....	11
3.7.1. Senaryolar .....	11
3.7.2. Use Case Modelleri .....	11
3.7.3. Statik Modeller .....	12
3.7.4. Dinamik modeller .....	15
3.8. Kullanıcı Arayüzü .....	19
<b>4. Sonuç .....</b>	<b>29</b>
<b>5. Referanslar .....</b>	<b>30</b>
<b>6. Terimler Listesi .....</b>	<b>31</b>
6.1. Tanımlar .....	31
6.2. Kısaltmalar .....	31

# 1. Giriş

Bu proje, söz konusu markanın duygusal içeriğini toplama, saklama ve analiz etme temellerini açıklamayı incelemektedir. Bu projedeki ana konular Yapay Zeka, Makine Öğrenimi, Derin Öğrenme ve Duygu Analizi alanındadır.

Bu projenin temel amacı, son kullanıcılar için toplanan verilerden duygusal bilgileri bulmak için etkili bir araç geliştirmek ve sağlamaktır.

Bu projenin ilk yarısında, web sitelerinden duygusal verilerin toplanması ve saklanması incelenecektir. Duygusal verilerin nasıl toplanacağını açıklayarak başlayacak . Bu sorun için mümkün olan tek seçenek, gerekli kütüphaneleri kullanarak duygusal verileri kazımdır. Bu iş için uygun görülen kütüphane BeautifulSoup'tur . Bu kütüphane hakkında gerekli bilgiler ilerleyen bölümlerde anlatılacaktır. Bundan sonra, bu verilerin daha sonra kullanmak üzere nasıl saklanacağını anlamaya gider.

Bu projenin ikinci yarısı, toplanan ve saklanan duygusal verilerin duygusal analizine odaklanıyor. Burada, toplanan verilerden duygusal bilgileri çıkarmak için kullanılan ortak ölçüler ve algoritmalar üzerinde durulmuştur. Ölçüler ve algoritmalar Neural Network, Naïve Bayes, Dictionary ve Word Cloud olarak belirlenmiştir. Bu yöntemler hakkında da gerekli açıklamalar ileriki bölümlerde yapılacaktır.

Ve Proje, görsel grafiklerle duygu analizinin gösterilmesiyle sona erer. Bu grafikler ile son kullanıcılar, yoğun görselleştirmeler yoluyla toplanan verilerin duygularını anlayabilecekler.

## 1.1. Projenin Amacı

Bu projenin temel amacı, son kullanıcılara toplumun bir sarf malzemesi, hizmet veya marka hakkındaki düşünceleri hakkında kısa ve duygusal bilgiler vermektir.

Bir markanın satışında müşteri memnuniyeti oldukça önemli bir faktördür. Tüketicilerin bir ürün, hizmet veya marka hakkında olumlu düşünceleri varsa bu, o markanın imajının oldukça iyi olduğu anlamına gelir. Ve eğer markanın kamuoyu iyiyse bu, bu şirketin büyük başarılar elde edebileceği ve daha fazla kar edebileceği anlamına gelir.

Günümüzde insanlar toplu olarak düşünce ve fikirlerini paylaşabilmektedirler. Ve web, bu tür duygusal veriler için en önemli bilgi kaynağıdır. Herhangi bir görüş veya fikir çok kısa sürede milyonlarca kişiye ulaşabilir. Bu kurumlar daha fazla kar elde etmek istiyor ve toplumun kendi hizmetleri, ürünleri veya imajları hakkında ne düşündüğünü önemsiyorsa, duygusal verileri de toplamayı ve analiz etmeyi düşünmelidir.

İş dünyasında, yukarıda belirtilen bu tür bir sorun için belirli bir şirket veya ürün hakkındaki düşünceleri analiz etmek için bazı sistemler vardır. Ancak bu projede her türlü web sitesinden veri toplayan bir sistem sağlanması hedeflenmiş ve bu verileri farklı duygu analizi yöntemleri kullanarak analiz eden bir program tasarlanmıştır.

## 1.2. Projenin Kapsamı, Çıktıları ve Tablolar

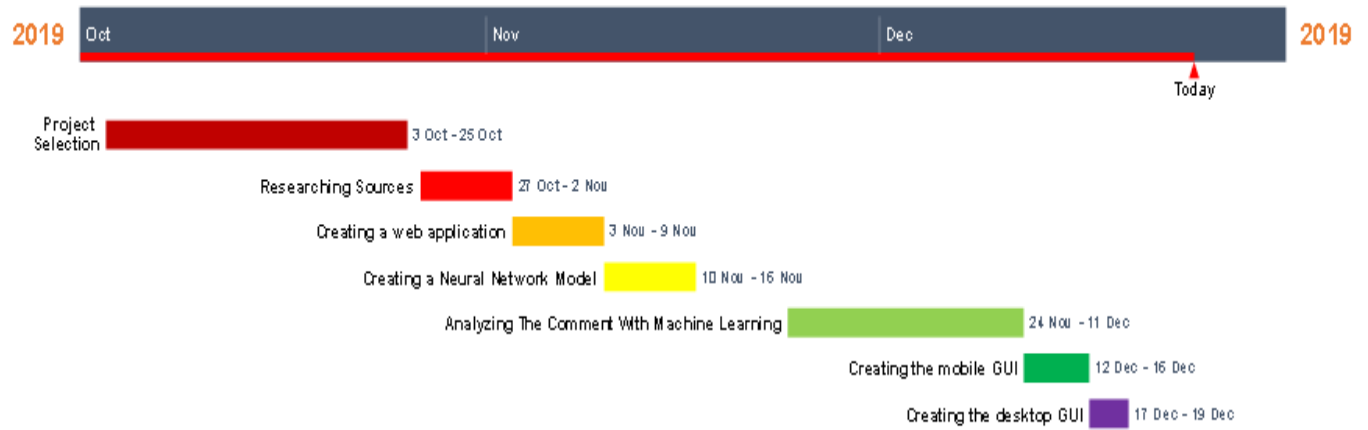
Projenin kapsamı, web sitelerinden duygusal bilgileri toplayan bir web kazıyıcıdan oluşan bir yazılım geliştirmek ve kullanıcılar için önemli grafikleri görüntülemek üzere çıkarılan verilerden anlamlı duygusal grafikler çıkaran bir analizör oluşturmaktır.

Bu proje, son kullanıcılar için bir yazılım programı ve bir yan bileşen için bir rapor üretecektir. Yazılım programı, belirli bir marka adıyla dünyanın her yerinden son yorumları tarama yeteneğine sahip bir kazıyıcı olacak. Ardından, verileri analiz edebilecek ve sonuçları bilgilendirici görselleştirmelerle gösterebilecektir.

Proje, 25.10.2019 tarihinde proje başlıklarının duyurulması ile başlamıştır. Bu proje raporunun önceki versiyonunda araştırmaya odaklanılmıştı. Projenin ikinci bölümünde ise projenin uygulanmasına odaklanıldı.

Project Selection	■ T	03/10/2019	25/10/2019	17 days
Researching Sources	■ T	27/10/2019	02/11/2019	5 days
Creating a web application	■ T	03/11/2019	09/11/2019	5 days
Creating a Neural Network Model	■ T	10/11/2019	16/11/2019	5 days
Analyzing The Comment With Machine Learning	■ T	24/11/2019	11/12/2019	13 days
Creating the mobile GUI	■ T	12/12/2019	16/12/2019	3 days
Creating the desktop GUI	■ T	17/12/2019	19/12/2019	3 days

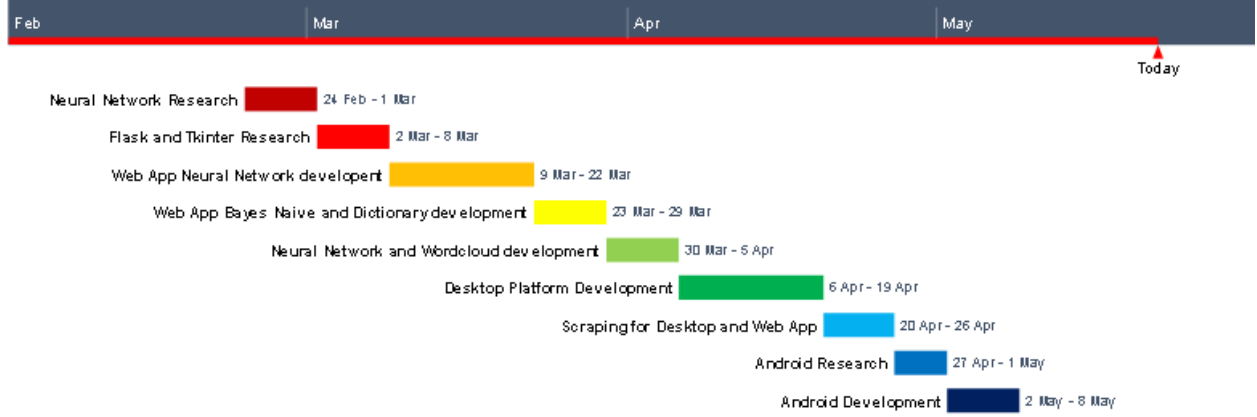
Şekil 1.2. 1- Proje Bölüm I Zaman Çizelgesi.



Şekil 1.2.2- Proje Bölüm I Zaman Çizelgesi.

Neural Network Research	■ T ▾	24/02/2020	01/03/2020	5 days
Flask and Tkinter Research	■ T ▾	02/03/2020	08/03/2020	5 days
Web App Neural Network developent	■ T ▾	09/03/2020	22/03/2020	10 days
Web App Bayes Naive and Dictionary development	■ T ▾	23/03/2020	29/03/2020	5 days
Neural Network and Wordcloud development	■ T ▾	30/03/2020	05/04/2020	5 days
Desktop Platform Development	■ T ▾	06/04/2020	19/04/2020	10 days
Scraping for Desktop and Web App	■ T ▾	20/04/2020	26/04/2020	5 days
Android Research	■ T ▾	27/04/2020	01/05/2020	5 days
Android Development	■ T ▾	02/05/2020	08/05/2020	5 days

Şekil 1.2. 3- Proje Kısım II Zaman Çizelgesi



Şekil 1.2.4- Proje II. Kısım Zaman Çizelgesi.

Projenin geliştirilmesi için zaman çizelgeleri yukarıda belirtilmiştir. ( Şekil 1.2.1), (Şekil 1.2.2), (Şekil 1.2.3), (Şekil 1.2.4). İlk iki rakam önceki bölümde tamamlanan işleri gösterir. Diğer iki grafik yapıları gösteriyor.

Bu projede hem Windows hem de Linux işletim sistemi kullanılmaktadır . Ana programlama dili, makine öğrenmesi için oluşturulmuş kullanışlı kütüphaneleri nedeniyle Python olarak belirlenmiştir. Bir web uygulama geliştirme aracı olarak Flask çerçevesi, masaüstü GUI için Tkinter kullanıldı. Mobil uygulama geliştirmek için ise Android Studio kullanılmaktadır. Analiz ve görselleştirmeler Matplotlib kütüphanesi ile gerçekleştirilmiştir .

Günümüzde birçok sosyal ağ, API anahtarlarını kolay kolay sağlamıyor. Ve diğer birçok API anahtarının kısıtlamaları vardır. Bu nedenle gerekli duygusal veriler python kütüphaneleri kullanılarak elde edilmiştir. BeautifulSoup kütüphanesi, web sitelerinden duygusal bilgileri çıkarmak için kullanılır. Ana dezavantaj, bu kütüphanenin web sitelerinden yalnızca statik içerik çıkarabilmesidir. Dinamik olarak yüklenen içeriği çıkaramaz.

### 1.3. Projenin Hedefleri ve Başarı Kriterleri

Projenin hedefleri ve başarı kriterleri şunlardır:

Aşağıdakileri yapabilen bir web ve masaüstü uygulaması sağlayın:

- Sosyal ağdan veri ayıklayın.
- Seçilen verilerin duygusal analizini yapın.
- Son kullanıcıya duygusal grafikler gösterin.

Aşağıdakileri yapabilen bir mobil uygulama sağlayın:

- Seçilen verilerin duygusal analizini yapın.
- Son kullanıcıya duygusal grafikler gösterin.

### 1.4. Proje Üyelerinin Görev ve Sorumlulukları

Tablo 1.4.1 Proje üyeleri, sorumlulukları ve görevleri

Proje Üyesi	Sorumluluklar ve Görevler
-------------	---------------------------

Doğacan Biçer	Araştırma, duygu analizi için gerekli verileri toplama, Toplanan verileri analiz etme, sonuçları görselleştirme, web, mobil ve masaüstü uygulaması oluşturma.
---------------	---

## 2. Yapılan Çalışma

Projeyi tamamlamak için iki şeyin gerçekleştirilmesi gerekir. Biri duygusal verileri kazımak. Diğeri, taranan duygusal verileri analiz etmektir. Duygusal verileri toplamak için iki yaklaşım dikkate alınmıştır. İlk yöntem, sosyal medya sitelerinin sağladığı API'leri kullanmaktır. "API" olarak da bilinen Uygulama Programlama Arayüzü, sosyal sistemlerin diğeri uygulama ve web sitelerinin kendi sosyal medya bilgilerini almasına ve kullanıcının web siteleri veya uygulamaları ile koordineli çalışmasına izin veren ilkeleri ifade eder. Sosyal ağın kendisi bu ilkeleri verdiği için, API kullanıcıları, Uygulama Programlama Arayüzünün adil bir model olmasını açıkça bekleyebilirler. API kullanıcısı tarafından gerçekleştirilebilecek kısıtlı nitelikler, Çağrı veya işlemler veriyorlar ve Uygulama Programlama Arayüzünü kullanmadan önce yazılım geliştiricinin bir API Anahtarı üretmesini gerektiriyorlar.[1]

İkinci yöntem, kazıma kitaplıklarını kullanmaktır. Web bilgi çıkarma olarak da bilinen web kazıma, içeriği çevrimiçi bir kaynaktan kurtarma veya kazıma yöntemidir. Dünya çapındaki web'in görünüşte sonsuz vahşi doğasından milyonlarca veya milyarlarca bilgiyi yakalamak için kurnazca robotlaştırmayı kurnazca yapan, fiziksel olarak ayıklayıcı bilgilerin sıradan ve bıktırıcı yönetimi gibi değil [2]. API kısıtlamaları nedeniyle bu projede bu seçenek kullanılmıştır ve scraping kütüphanesi BeautifulSoup olarak seçilmiştir .

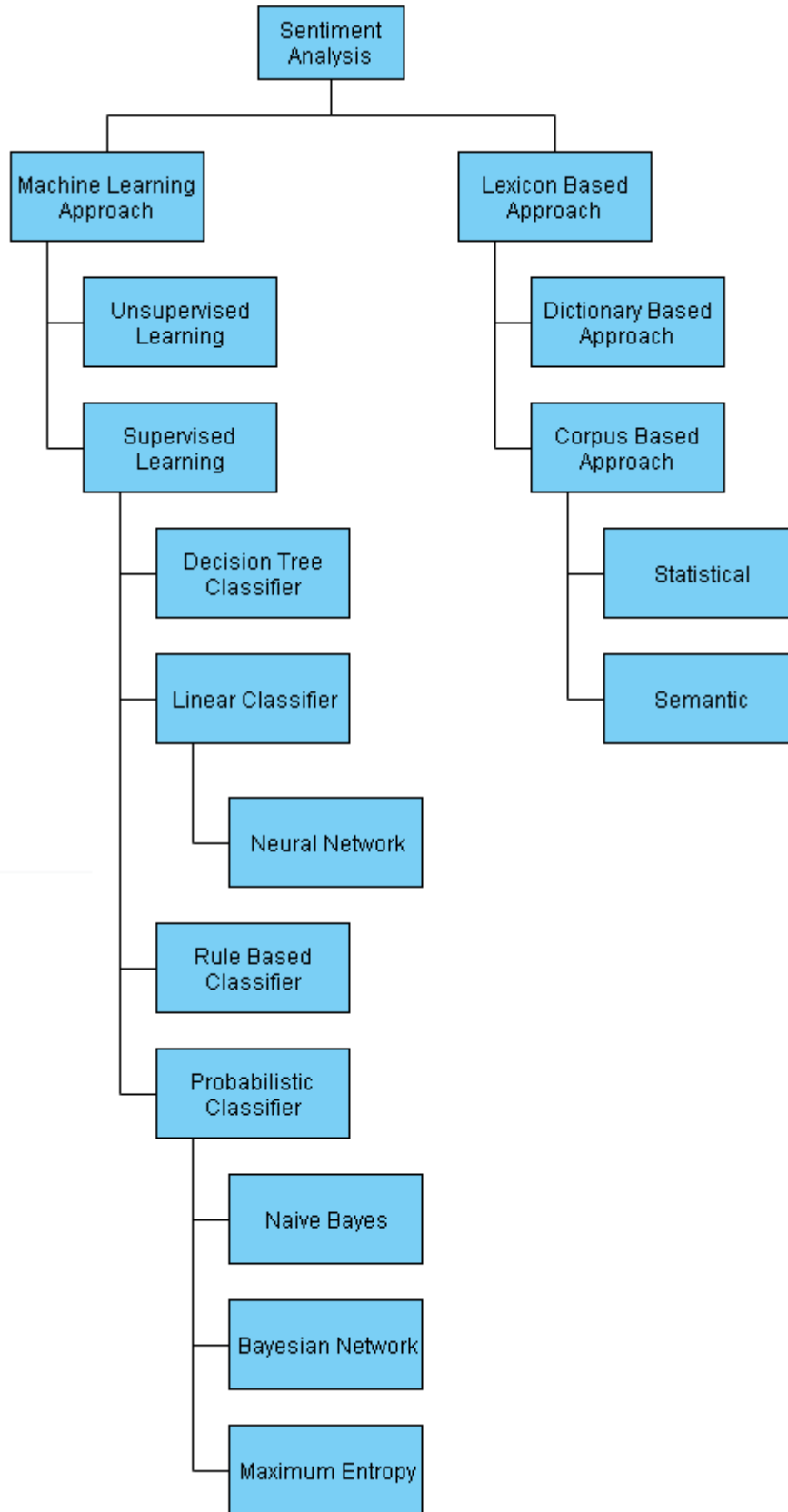
Gerçekleştirilmesi gereken diğeri şey, toplanan verilerin analiz edilmesidir. Duyguyu analiz ederken temel olarak iki ana yöntem vardır. Bir yaklaşım, Makine Öğrenimi yaklaşımıdır. Diğeri Lexicon tabanlı yaklaşımdır. Bunların da alt dalları vardır.

Makine Öğrenimi Yaklaşımı, içeriği tanıyabilen bir sınıflandırıcı oluşturmak için makine öğrenimi stratejilerini ve farklı nitelikleri kullanır. Bu günlerde derin öğrenme stratejileri, Tekrarlayan Sinir Ağları gibi bilgi öğrenme temsillerine uydukları için iyi biliniyor.[3]

Sözlüğe Dayalı strateji, belirli bir maddenin ortak değerlendirme puanını seçmek için ekstremiter puanıyla açıklanan bir dizi kelime kullanır. Bu prosedürün en yararlı kaynağı, herhangi bir hazırlık bilgisi gerektirmemesi, dezavantajı ise varsayım sözlüklerinde çok sayıda kelime ve ifadenin yer almamasıdır.[4]

Basitçe duygu analizi aşağıdaki grafik ile ifade edilebilir. ( Şekil 2.1 )





Şekil 2.1- Duygu Analizi Türleri.

## 2.1. Makine Öğrenmesi

BeautifulSoup , bilgileri ayrıştırmak, yani bilgileri HTML veya XML dosyalarından ayıklamak için programlanmış Python paketidir.

BeautifulSoup can bilgileri ayrıştırmak için kullanıldığından ve html ve xml dosyalarını kendisi alamadığından, Request adlı kitaplık ile düzenli olarak kullanılmaktadır. Bu gibi durumlarda, İstek kitaplığı siteden internet sayfasını almasını istemek için HTTP yapacak ve sayfa döndürüldüğünde, HTML sayfasındaki hedef bilgileri ayrıştırmak için BeautifulSoup paketi kullanılabilir.

BeautifulSoup'u kullanmanın en büyük odak noktalarından biri, zahmetsiz olması ve kazıma sırasında bilgileri ayrıştırmanın yinelenen kısımlarından birkaçını bilgisayar haline getirme kapasitesidir. Kullanıcı, az sayıda kod kullanırken, ayrıştırılmış bir arşivin tamamını keşfetmek ve ihtiyaç duydukları bilgilerin tüm fırsatlarını keşfetmek veya olağanüstü karakterler gibi kodlamaları doğal olarak tanımlamak için BeautifulSoup'u tasarlayabilecektir . [5]

## 2.2. Sinir Ağı

Istm düğümleri arasındaki ilişkilerin dünyevi bir grupta boyunca koordineli bir çizelge şekillendirdiği bir sinir ağı türüdür [6].

bir düzenlemenin her bileşeni için aynı atamayı yapabildikleri ve sonuç geçmiş sonuçlarla ilişkili olduğu için tekrarlayan olarak adlandırılır. Başka bir deyişle, Tekrarlayan Sinir Ağlarının, henüz hesaplanmış olana yakın verileri yakalayan kısa bir belleğe sahip olduğunu varsayabiliriz. Hipotezde, Tekrarlayan Sinir Ağları, verileri iddialı uzun düzenlemelerde kullanabilir, ancak gerçekte, birçok adım geriye bakamazlar. [7]

İleri beslemeli çalışan yapılarda, gelen bilgiler sadece ileri doğru işlenir. Bu tür yapılarda, bir çıkış değeri elde etmek için giriş verileri ağdan geçirilir. Ortaya çıkan çıkış değeri, bir hata elde etmek için doğru değerlerle karşılaştırılır. Ağdaki ağırlık değerleri hataya göre değişir, bu şekilde oluşturulan model ile en doğru sonucu üretebilir.

Bir ileri beslemeli ağı eğitirken, hata yeterince azaltılmalıdır. Böylece nöronlara verilen ağırlıklar yenilenir ve girdiye uygun çıktıyı verecek bir yapı oluşur. Tekrarlayan yapılar, çıktılarını bir sonraki süreçte girdi olarak kullandıkları için ileri beslemeli yapılardan ayrılırlar. Başka bir deyişle, bu ağlardaki girdiler, mevcut ve önceki bilgileri birleştirerek çıktı üretir.

## 2.3. Bayes Algortiması

Naive Bayes Sınıflandırıcılar, metin sınıflandırma problemlerini çözerken oldukça kullanışlıdır. Metin sınıflandırmalarının amacı, belgeleri bir veya daha fazla sınıfa ayırmaktır. Naive Bayes kullanımının bazı uygulamaları Öneri Sistemleri, Spam Filtreleme, Çoklu Sınıf veya Gerçek Zamanlı Tahmindir.

Naive Bayes sınıflandırıcısının başka bir örneği, amacı yazarın belirli bir konu, öge, fayda vb. hakkında bakış açısına karar vermek olan Duygu Analizidir. [8]

## 2.4. Sözlük Tabanlı Yöntem

Sözlük Tabanlı yöntem şöyle çalışır: Öncelikle olumlu ve olumsuz kelimeler toplanır ve saklanır. Daha sonra toplanan verilerdeki her kelime bu kelime çantasıyla karşılaştırılır. Toplanan verilerde negatif kelimelerden daha fazla pozitif kelime varsa, bu yorumun pozitif olduğu anlamına gelir. Aynı şekilde, yorumda olumlu kelimelerden daha fazla olumsuz kelime varsa, bu yorum olumsuz olur .

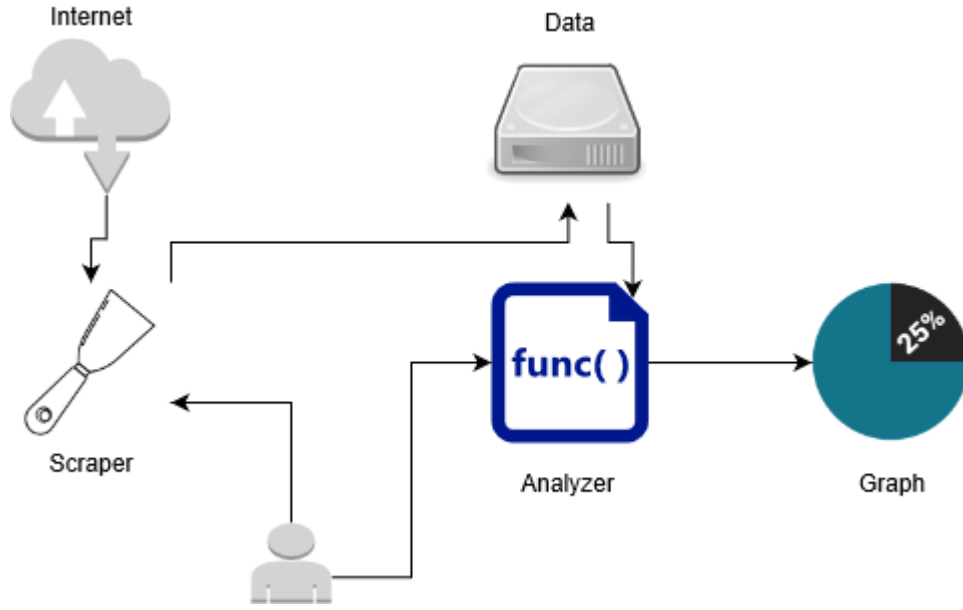
Bu yöntem kullanılırken olumsuzluk eki de dikkate alınır. Örneğin, olumsuz kelime olarak kabul edilen “beğenme” bile olumlu kelimelere aittir. Çünkü önceki kelime olumsuzdur.

## 3. Önerilen Sistem

### 3.1. Genel Bakış

Web uygulama ve masaüstü platformları için sistemin 4 ana süreci vardır. Süreçler, toplanan verileri Kazıma, Depolama, Analiz ve Görselleştirme'dir. Web gönderilerini indirmek için arka planda bir kazıyıcı çalışır. Verileri indirdikten sonra, kullanıcı bu veri sayfasını analizörlere ekleyebilir. Son olarak, analizör verilere çeşitli teknikler uygular ve elde edilen görüntüleri kullanıcıya gösterir. Sistem mimarisi aşağıda gösterilmiştir.

Mobil uygulama kısmı için, kullanıcı verileri yalnızca sözlük yöntemiyle analiz edebilir. Bu platform için sistem mimarisi aşağıda gösterilmiştir.



Şekil 3.1- Önerilen sistem mimarisi

İlk adım, analizörler için gerekli verilerin toplanmasıdır. Gerekli veriler, web sitelerinde yayınlanan gönderiler olacaktır .

İkinci adım, verilerin saklanmasıdır. Kullanıcının bir seçeneği vardır, onu yerel bir sürücüde saklayabilir.

Son adım, verilerin analiz edilmesidir. Verilerin analizinde Naïve-Bayes, Neural Network, Dictionary ve Word Cloud teknikleri kullanılacaktır. Son olarak masaüstü ve web platformlarında matplotlib yardımıyla, mobil platformda MPAndroidChart grafik kütüphanesi ile sonuç kullanıcıya gösterilecektir .

### 3.2. Analiz ve Görselleştirme Araçları

Web uygulaması geliştirilirken Flask çerçevesi kullanılır. Flask, Python'da yazılmış bir mikro WSGI web çerçevesidir. Flask, başlangıç sürecini hızlı ve net hale getirmek için geliştirilmiştir. Flask çerçevesini kullanarak, kullanıcılar karmaşık programlar da oluşturabilirler [9].

Masaüstü tarafında grafik arayüz olarak Tkinter kullanılmıştır. Tkinter , Python programlama dilinde yazılmış bir grafik kullanıcı arayüzüdür. Tkinter , Unix sistemlerinde, Windows veya Linux'ta kullanılabilir [10]. Tkinter'ı bu projede kullanmanın asıl amacı, platformlar arası özelliğinden kaynaklanmaktadır. Hem Windows hem de Linux işletim sistemlerini çalıştırır.

Mobil tarafta programlama dili olarak java, geliştirme aracı olarak android studio kullanılacaktır. Tüm platformların aynı veri tabanını paylaşması planlanmaktadır. Bu sayede web uygulamasından veri tabanına yüklenen bir veri sayfası mobil platformda analiz edilebilmektedir. Web ve masaüstü uygulamasında, kullanıcıya grafikleri göstermek için python'un matplotlib kütüphanesi kullanılacaktır. MPAndroidChart kütüphanesi ve mobil tarafta grafik görünüm kütüphanesi olarak belirlenmiştir.

Bu uygulama, topluluğun markaları hakkında ne düşündüğünü öğrenmek isteyen kurumlara hizmet edecek. Programın asıl amacı toplulukta markaların imajlarını göstermek olsa da girilen anahtar kelimeye göre ürünler, kişiler ve olaylar hakkında duygusal bilgiler de verebilmektedir.

### 3.3. Önerilen Sistem Mimarisi

Web uygulaması ve Masaüstü platformları için Hurda bölümünde, kullanıcılar bir web sitesindeki metni hurdaya çıkarabilir. Bu bölüm, BeautifulSoup paketini kullanır. Web sitesi link kısmında, Kullanıcılar web sitesinin linkini gireceklerdir. Ardından, ilgilendikleri metni içeren etiketi gireceklerdir. Ve eğer sınıf varsa etiketinin sınıfını yazabilirler. Eğer tag ve class kısmına bir şey yazmazlarsa, program tag kısmına “p” yazdığını varsayar ve bu tagın classı yoktur. Arka uçta, BeautifulSoup ve Request Libraries metinleri indirir.

Ancak bu bölüm, dinamik olarak yüklenen içeriği hurdaya çıkarmaz. Çıktı csv dosyası olarak görünecek. Ve program otomatik olarak noktalama işaretlerini kaldırır, sayıları ve kelimeleri düşürür.

Kullanıcılar ayrıca bir sinir ağı modeli oluşturabilir. Bir csv seçtiklerinde ve sinir ağı modeli bilgisini girdiklerinde, Keras kitaplığı bir sinir ağı modeli oluşturur. Arka uçta TensorFlow kitaplığını kullanır. CSV'de "pozitif", "negatif" veya "nötr" gibi metinler ve bunların duyguları olmalıdır. Ve model bittiğinde, Kullanıcı modeli indirebilir veya kaydedebilir. Model h5 dosyası olarak kaydedilecek. Artık Kullanıcılar bu modeli istedikleri zaman kullanabilirler.

Analiz bölümünde, Kullanıcı diğerleri arasından bir duyarlılık analiz yöntemi seçebilir. Neural Network kullanarak verileri analiz etmek istiyorlarsa, bir modele sahip olmaları gerekir. “Model Oluştur” bölümünden sinir ağı modeli oluşturulabilir. Model seçildikten sonra, artık yorumların veya metinlerin bulunduğu bir csv seçmeleri gerekiyor. CSV, metinlerden oluşmalıdır. Ve program csv'yi analiz eder ve ortaya çıkan grafiği kullanıcıya gösterir.

Kullanıcılar ayrıca Naïve Bayes Yaklaşımı kullanarak metinleri analiz etmek isteyebilirler, bu yöntem için bir model Dosyasına gerek yoktur. Ancak kullanıcıların yorumları ve "olumlu" veya "olumsuz" gibi duygusal bilgilerini içeren bir csv dosyasına ihtiyacı vardır. Yöntem bu dosyayı eğitim dosyası olarak kullanacak. Ve metin içeren başka bir csv'ye ihtiyaçları var. Ve program csv'yi analiz eder ve ortaya çıkan grafiği kullanıcıya gösterir.

Dictionary yöntemi adı verilen başka bir yöntem daha vardır. Bu yaklaşımda pozitif ve negatif csv vardır. Program, seçilen csv'deki her kelimeyi karşılaştırır ve o cümleyi puanlar. CSV'deki tüm kelimeleri karşılaştırır ve ortaya çıkan grafiği gösterir. Bunu yaparken kelimedenden önce gelen olumsuzluk ekini de dikkate alır.

### 3.4. Fonksiyonel Gereksinimler

Proje, zamana kadar aşağıdaki gereksinimleri karşılayabilmelidir.

- Proje, web sitelerinden veri toplayan bir kazıyıcı oluşturur.
  - Web sitelerinde yapılan yorumlar.
- Daha sonra web sitelerinde yayımlanan metinlerden bir sonuca ulaşabilmek için bu verileri analiz eder.
  - Bir Sinir Ağı
  - Naive Bayes Metin Sınıflandırması
  - Sözlük Yöntemi
  - Kelime Bulutu
- Nihayet, BT verir A görsel temsil ile ilgili analiz için insan ile analiz itibaren farklı açılar.

### 3.5. Fonksiyonel Olmayan Gereksinimler

#### 3.5.1. Kalite Gereksinimleri

##### 3.5.1.1 Özel Kalite Gereksinimleri

###### a) Güvenilirlik

Web uygulama platformu sunucuda çalıştığı için bu platformun güvenilirliği sunucunun kullanılabilirliğine bağlıdır. Ayrıca yorumların silinmesi internet bağlantısının olup olmamasına bağlıdır.

Program birçok kez test edilmiş ve sistemin güvenilir sonuçlar verdiği kanıtlanmıştır.

###### b) Kullanılabilirlik

Bu programlar çok kullanıcı dostu değildir. Bu yazılımı kullanmak için kullanıcının bir şeyler bilmesi gerekir. Örneğin, yorumları hurdaya çıkarmak için, kullanıcının web sitesinin kaynak kodunu görmesi ve metnin etiketini ve sınıfını belirlemesi gerekir. Ancak kullanıcı arayüzü karmaşık değildir. Her analiz yöntemi ayrı bölümlere ayrılmıştır. Ve tüm yöntemler ana sayfada mevcuttur.

Ayrıca, masaüstü platformunda bir sıralama var. Örneğin, ilk kullanıcı iki dosya seçmeli, ardından analiz et ve kaydet'e tıklayın.

Mobil platform oldukça kullanıcı dostudur. Kullanıcıların tek yapması gereken CSV'yi seçmektir. Program kullanıcıya grafiği otomatik olarak gösterir. Bu nedenle, web uygulaması ve Masaüstü platformlarında bir kullanım kılavuzuna ihtiyaç duyulacaktır.

###### c) Güvenlik

Web üzerinde çalışan birçok uygulamada olduğu gibi web uygulama platformunda da güvenlik açıkları bulunabilir. Fark edilen tüm güvenlik açıkları kaldırıldı.

###### d) Kullanılabilirlik

Web uygulaması sunucu üzerinde çalışıyorsa, web uygulamasının kullanılabilirliği sunucunun kullanılabilirliğine bağlıdır. Ve yorumları çöpe atmak için Masaüstü ve Web Uygulama Platformu için internet bağlantısı gereklidir. Diğer bölümler ve mobil platform için internet bağlantısı gerekli değildir.

###### e) Sürdürülebilirlik

Programın daha iyi anlaşılması ve daha iyi organize edilmesi için program fonksiyonlara ve sınıflara bölünmüştür. Yani nesne yönelimli programlama kullanılır. Bu sayede, bir şeyin değiştirilmesi gerekiyorsa, kolayca değiştirilebilir .

### 3.5.2. Performans Gereksinimleri

#### a) Zamanlama

Analizörlere yapılan yorum sayısı fazla ise, duyarlılık oranı hesaplaması daha uzun sürecektir. Yorumlardaki kelime sayısı da süreyi etkiler.

#### b) Hız

Programın hızı kullanıcının donanımına bağlıdır. Özellikle sinir ağı oluşturmak için kullanıcının donanımı oldukça önemlidir. Kullanıcının donanımı ne kadar iyiye, program o kadar hızlı çalışacaktır.

#### c) Hacim

Program fazla yer kaplamaz. Her üç platformun da 100 mb'den daha az alana ihtiyacı var.

#### d) Verim

Program, kullanıcıya duyarlılık grafiklerini gösterecek. Grafikler, yorumların olumlu, olumsuz ve nötr oranını gösterecektir. Ayrıca program, yorumları csv olarak hurdaya çıkarabilecek ve h5 dosyası olarak bir sinir ağı modeli oluşturabilecektir.

## 3.6. Gerçek Olmayan Gereksinimler

### 3.6.1 Fiziksel

Sistemin mobil sürümünün çalışması için bir android mobil cihaza ihtiyacı vardır. Masaüstü tarafı için, Sistem Linux veya Windows işletim sistemini çalıştırabilen bir platforma ihtiyaç duyar. Web platformunda, herhangi bir web tarayıcısını çalıştırabilen herhangi bir cihaz yeterli olacaktır.

### 3.6.2 Yasal

Kazıma yöntemiyle elde edilen verilerin analiz edilmesi ve satılması hukuki açıdan sorun teşkil edebilmektedir. Ayrıca sitelerden veri çekerken robots.txt dosyasına saygı gösterilmelidir. Örneğin, kazımayı yasaklayan web sitelerinden herhangi bir veri alınmamalıdır. Robots.txt dosyasında bundan bahsetmişlerdi.

### 3.6.3 Kültürel

İnsanların cinsiyeti, ırkı ve dinine göre çıkarımlarda bulunmak etik olmayacaktır. Kültürel etik için bu projede yorumcuların cinsiyet, din ve ırk ayrımı yapılmayacaktır . yorumcu bilinmeyecektir.

### 3.6.4 Çevre

Mobil platformda program sadece android işletim sisteminde çalışacaktır. Masaüstü ve dizüstü bilgisayarlarda, programın çalışması için Windows veya Linux işletim sistemi gerekir.

### 3.6.5 Tasarım ve Uygulama

Sistem, Python ve Java ortamları kullanılarak geliştirilmiştir. Sistem bu dillere ait kütüphaneler kullanılarak geliştirilmiştir.

### 3.6.6 Arayüz

Bu projede Python ve Java programlama dilleri kullanılmıştır. Bu nedenle, web ve masaüstü platformları, python ile uyumlu bir Grafik Kullanıcı Arayüzüne ihtiyaç duyar. Tkinter bu iş için kullanılır. Web uygulama platformunda Flask çerçevesi kullanılır. Her türlü web geliştirme kütüphanesi flask ile kullanılabilir.

Önyükeme çerçevesi, web uygulama platformunun kullanıcı arabirimini geliştirirken kullanılır. Mobil tarafta, proje kullanıcı arabirimi, android stüdyo geliştirme ortamına bağlıdır.

### 3.7. Sistemdeki Modeller

- senaryolar

**Senaryo: Web uygulaması ve masaüstü Platformları için bir markanın duygu puanlarının analizi.**

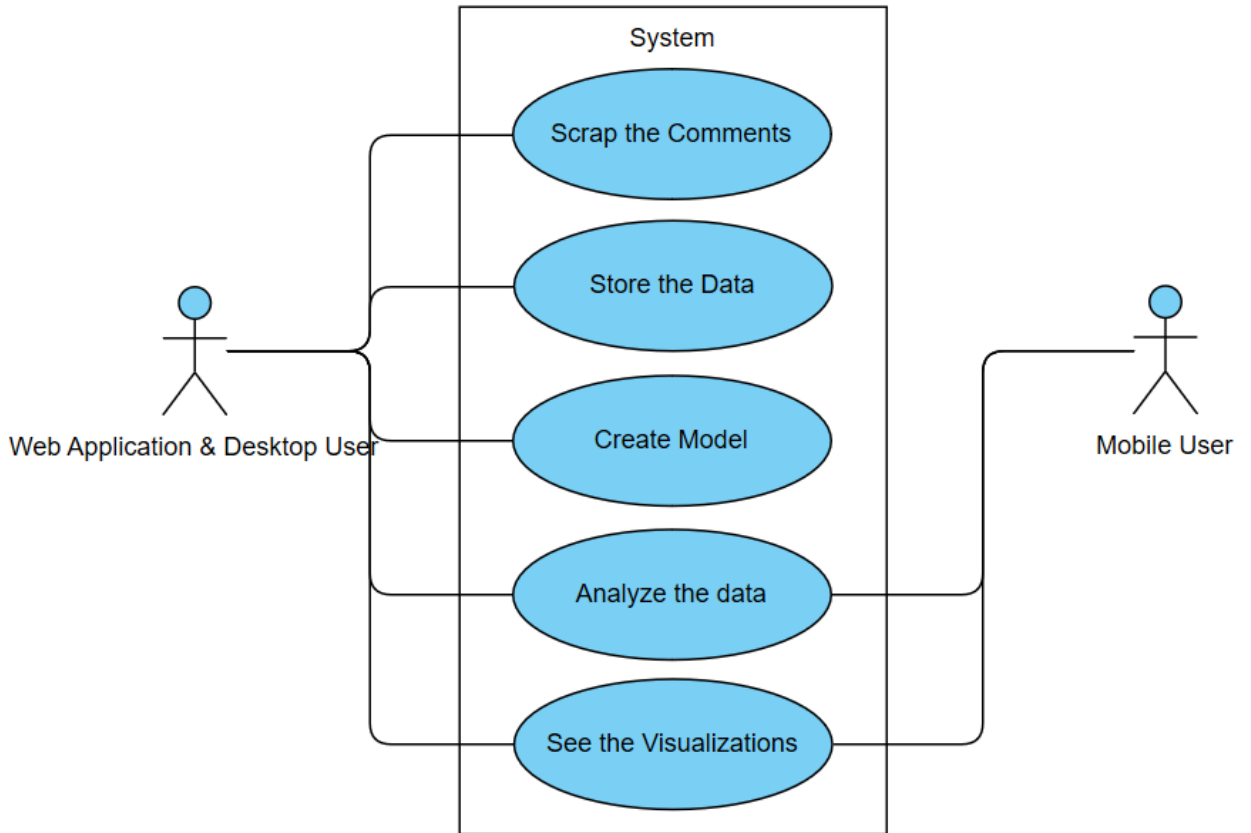
1. Kullanıcı web sitesi bilgilerini girer.
2. Kullanıcı bu verileri kaydeder.
3. Kullanıcı bu verileri analiz araçlarıyla analiz eder.
4. Kullanıcı analiz edilen grafikleri indirir.

**Senaryo: Web uygulaması ve masaüstü Platformları için bir markanın duygu puanlarının analizi.**

1. Kullanıcı verileri seçer.
2. Kullanıcı bu verileri sözlük yöntemini kullanarak analiz eder.

- Kullanım Örneği modeli

Programın ana kullanım durumları aşağıda gösterilmiştir .

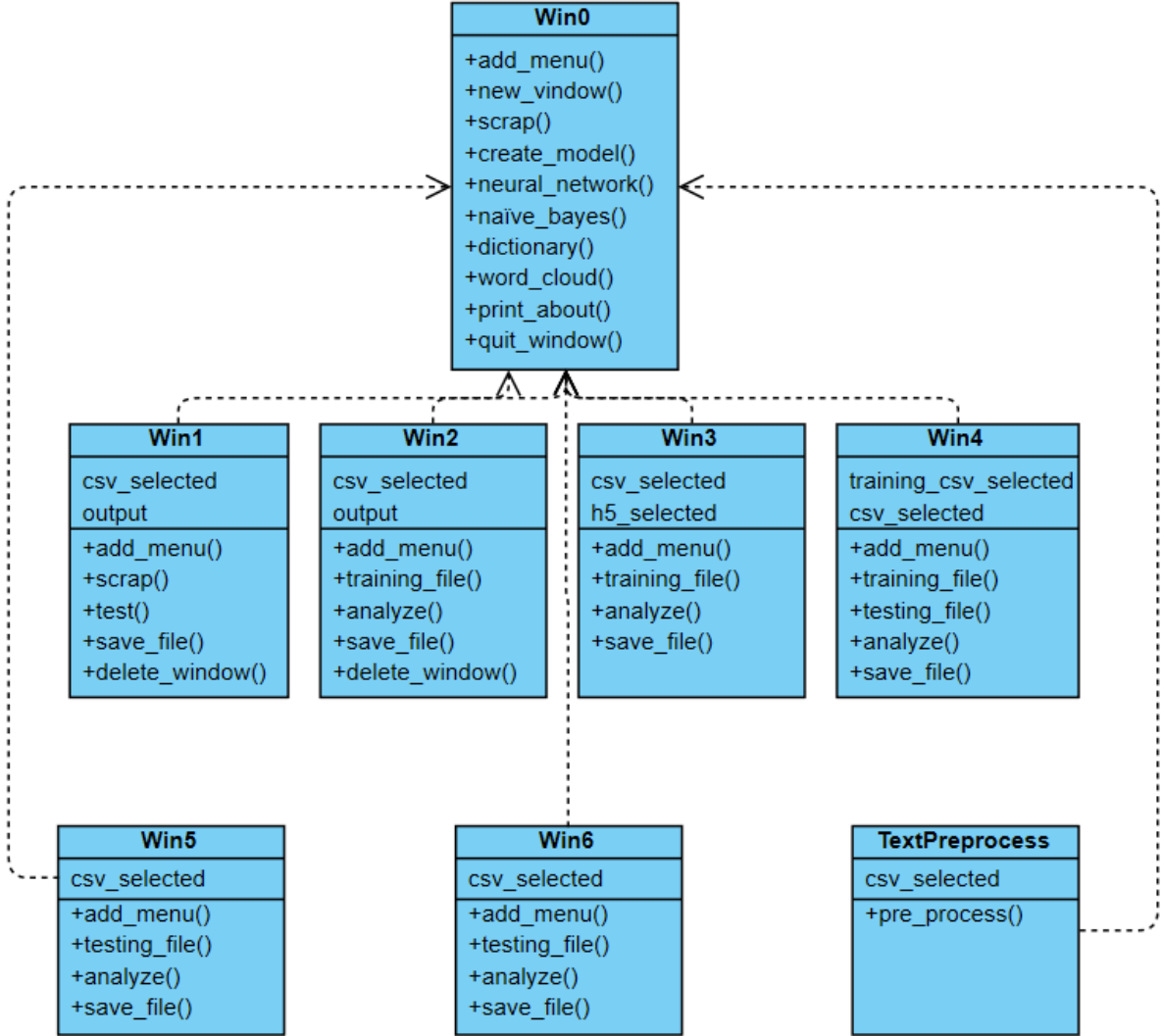


Şekil 3.7.1- Kullanıcının Kullanım Durumu Modeli.

- Statik modeller

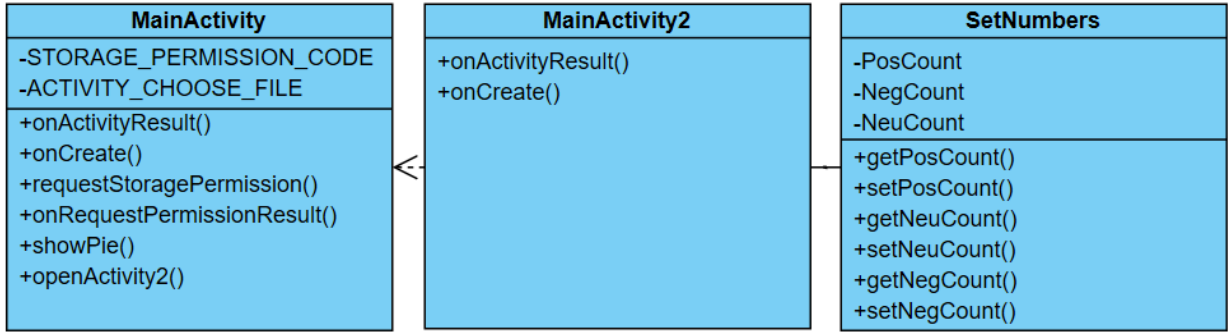
- Sınıf diyagramı

Programın sınıf diyagramı aşağıda gösterilmiştir. Aşağıdaki diyagramlar, masaüstü ve mobil uygulamanın sınıf diyagramlarını temsil etmektedir. Web Uygulama platformu için sınıf yoktur.



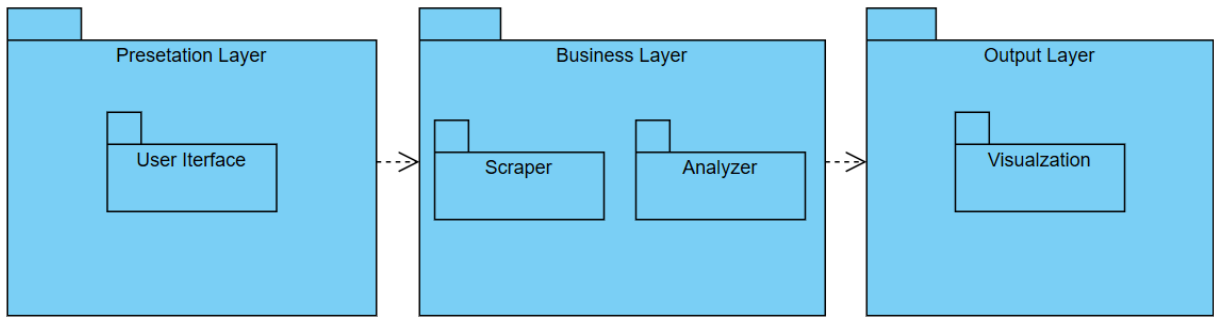
Şekil 3.7.2- Masaüstü Platformda Projenin Sınıf Diyagramı





Şekil 3.7.3 - Mobil Platformda Projenin Sınıf Diyagramı

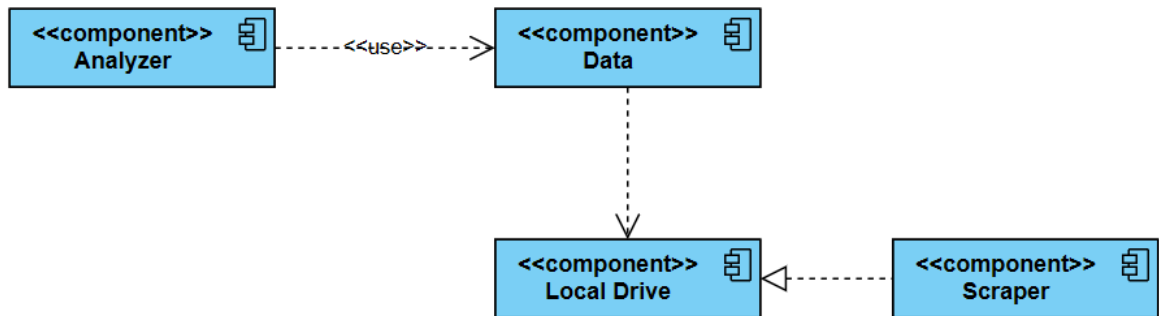
### • Paket Şeması



Şekil 3.7.4- Projenin Paket Şeması

### • Bileşen Şeması

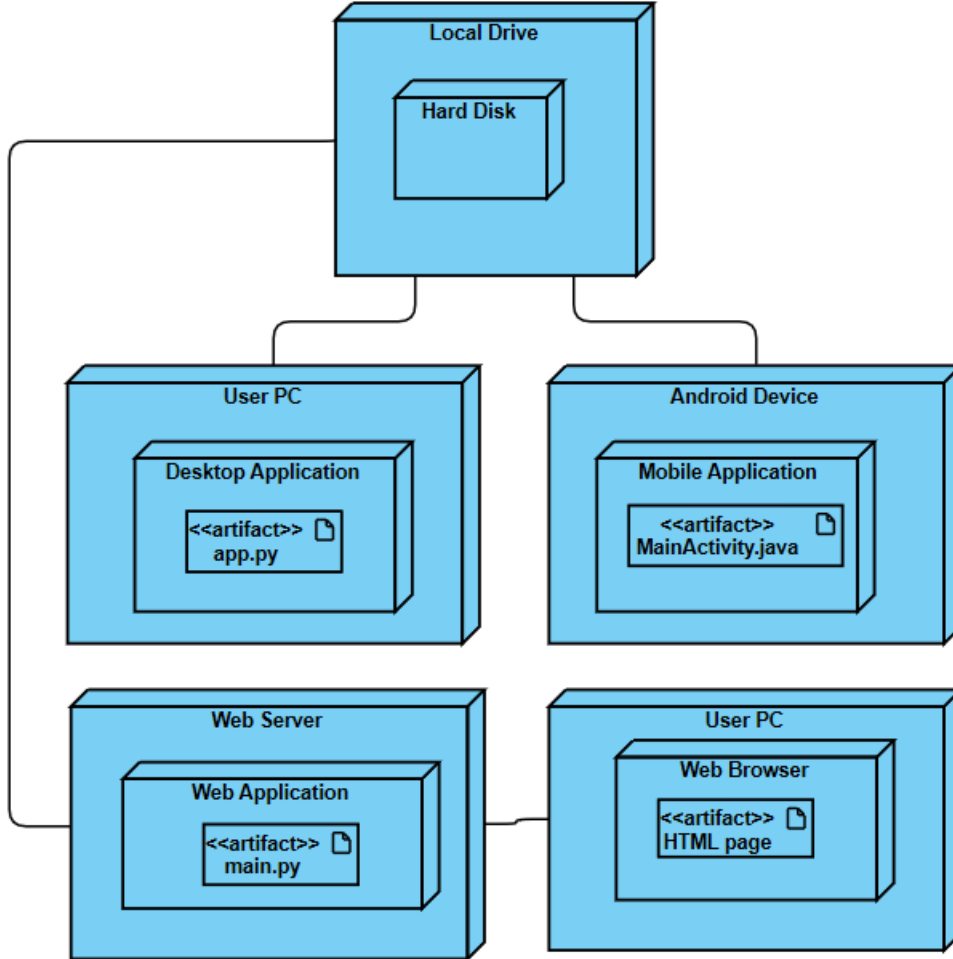
Programın bileşen şeması aşağıda gösterilmiştir.



Şekil 3.7.5- Projenin Bileşen Şeması

- **Dağıtım Şeması**

Programın Dağıtım şeması aşağıda gösterilmiştir. Yerel sürücü, platformdan platforma farklılık gösterebilir. Temel olarak, yorumların sistemde depolandığı yeri temsil eder.

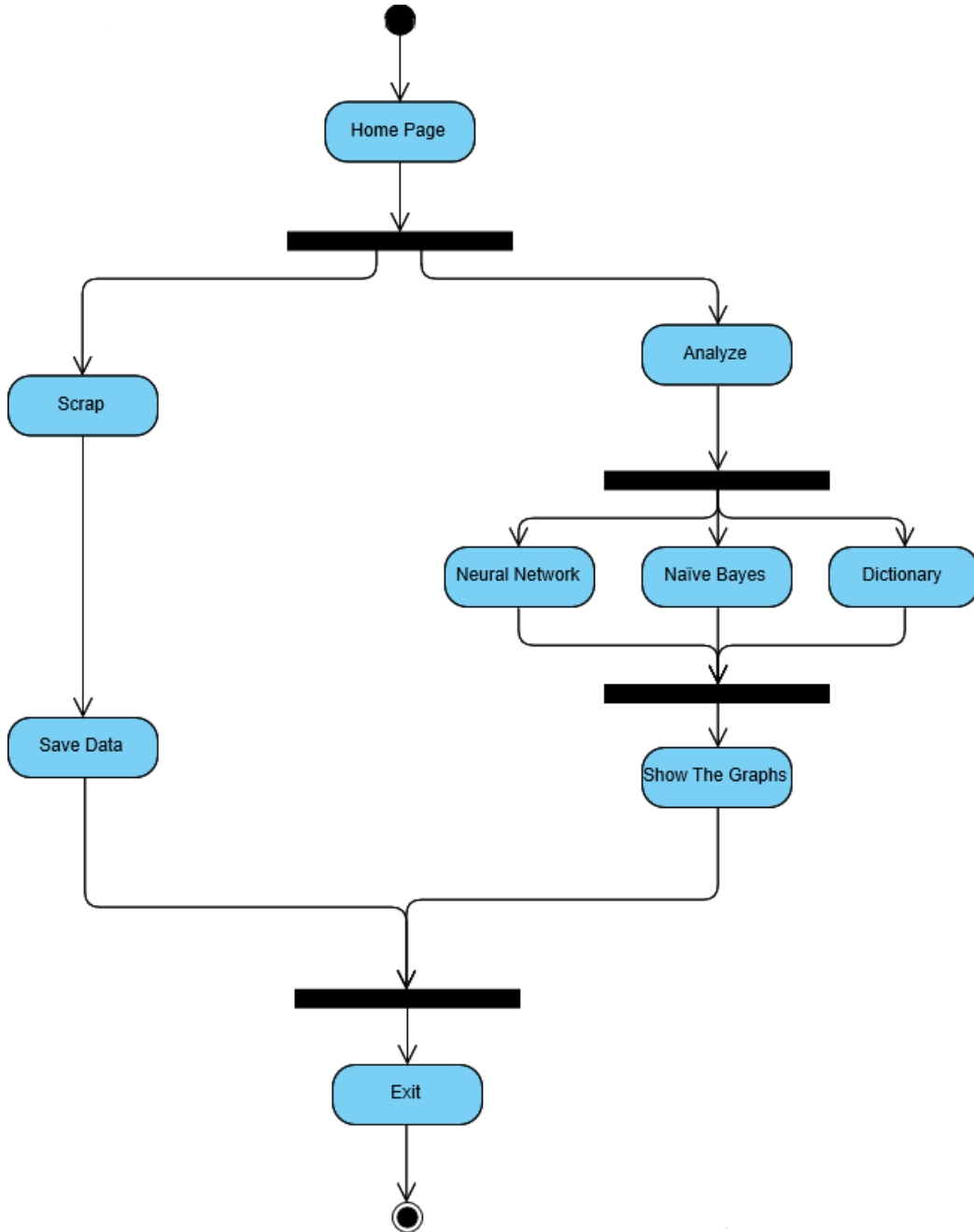


Şekil 3.7.6 - Projenin Yerleştirme Şeması

- Dinamik modeller

- Etkinlik şeması

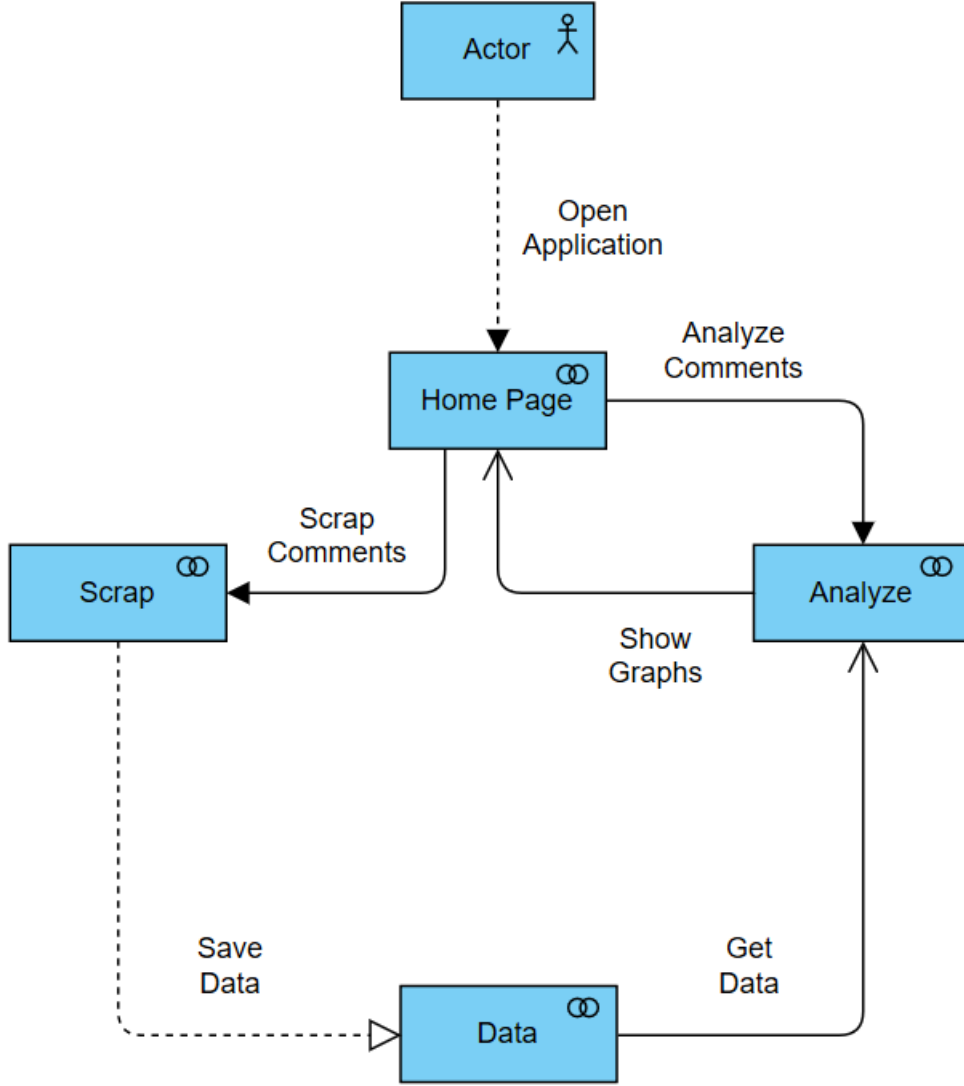
Programın aktivite diyagramı aşağıda gösterilmiştir. Aşağıdaki diyagram, masaüstü ve web uygulama platformu için aktivite diyagramını temsil eder. Mobil platformda sadece Ana Sayfa, Analiz Et, Sözlük ve Grafiği göster aktiviteleri bulunmaktadır.



Şekil 3.7.7- Projenin Faaliyet Şeması

## • İşbirliği Şeması

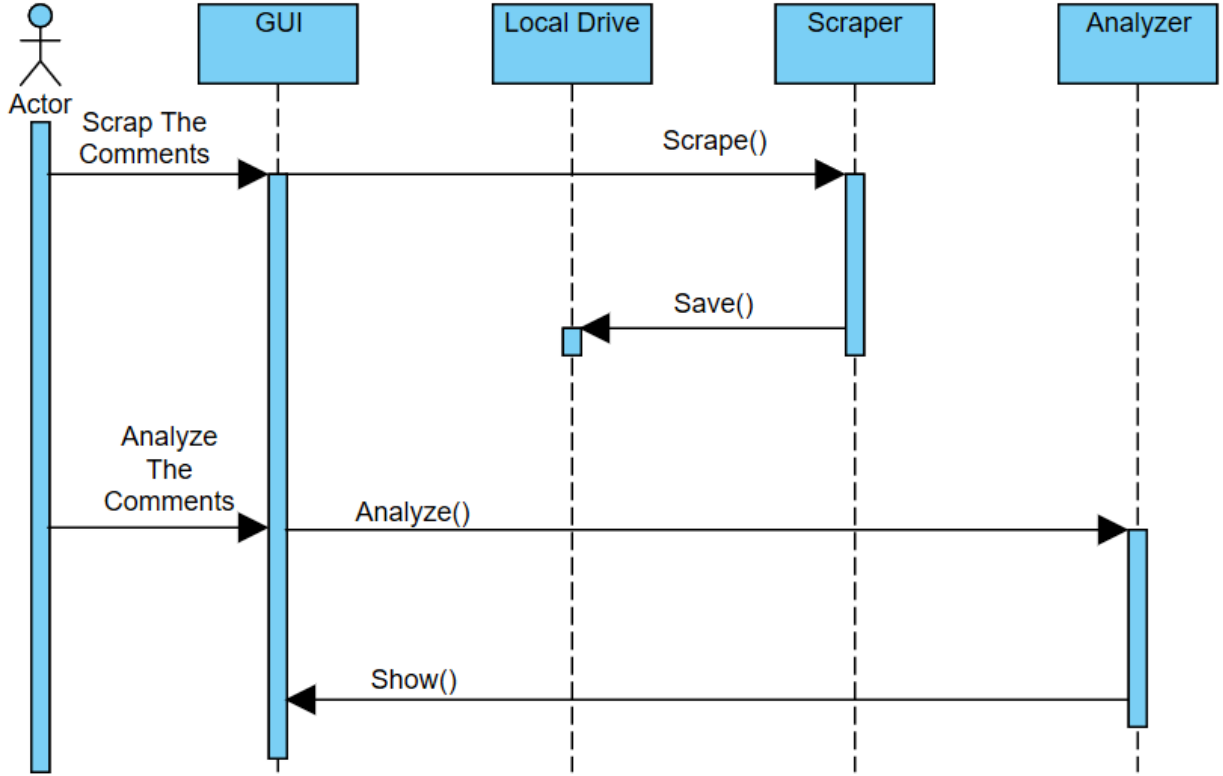
Programın işbirliği şeması aşağıdaki şemada gösterilmiştir. Aşağıdaki şema, Masaüstü ve web uygulama platformunun işbirliği şemasını göstermektedir. Mobil uygulama için sadece Ana Sayfa, Analiz ve veri işbirlikleri vardır.



Şekil 3.7.8- Projenin İşbirliği Şeması

- **Sıra Diyagramı**

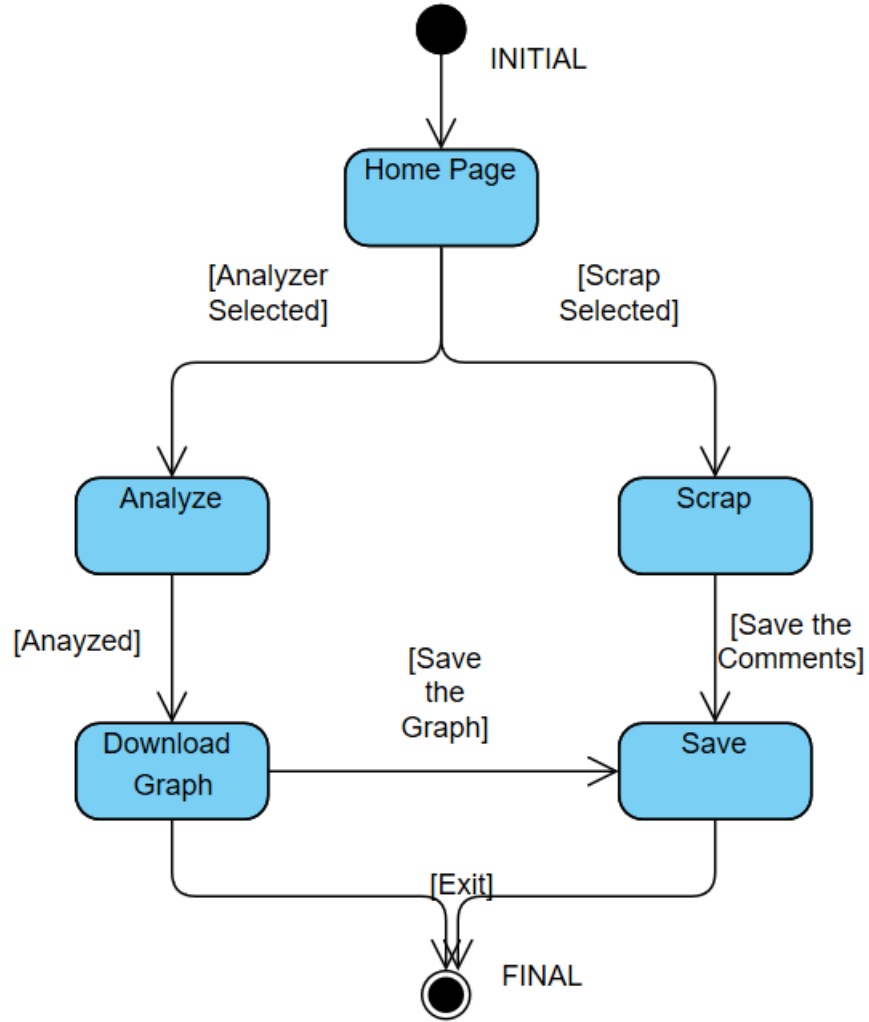
Programın sıralama şeması aşağıdaki şemada gösterilmiştir. Aşağıdaki şema, Masaüstü ve web uygulama platformunun sıralama şemasını göstermektedir. Mobil uygulama için sıyrıcı sıralaması yoktur.



Şekil 3.7.9- Projenin Sıra Şeması

- **Durum Grafiği Diyagramı**

Programın durum şeması diyagramı aşağıdaki şemada gösterilmiştir. Aşağıdaki şema, Masaüstü ve web uygulama platformunun durum şeması diyagramını göstermektedir. Mobil uygulama için sadece Ana Sayfa ve analiz durumları bulunmaktadır.



Şekil 3.7.10 - Durum Tablosu Proje Şeması

### 3.8. Kullanıcı arayüzü

Ana sayfada iki bölüm bulunmaktadır. Hurda kısmında, kullanıcı web sitesindeki yorumları hurdaya çıkarabilir veya bir sinir ağı modeli oluşturabilir. Kullanıcılar üretilen verileri kaydedebilir ve bu verileri daha sonra kullanmak üzere kullanabilirler. Sağ tarafta ise analiz kısmı bulunmaktadır. Bu bölümde, kullanıcılar verilerini analiz etmek için Neural Network, Naïve Bayes veya Dictionary Yöntemlerini kullanabilirler. Veya bir kelime bulutu çizebilirler.

Scrap	Analyze
<input type="button" value="Scrap"/>	<input type="button" value="Neural Network"/>
<input type="button" value="Create Model"/>	<input type="button" value="Naive Bayes"/>
	<input type="button" value="Dictionary"/>
	<input type="button" value="Wordcloud"/>

Şekil 3.8.1 -Web Uygulaması Kullanıcı Arayüzü

Aşağıdaki bölümde, kullanıcılar bir web sitesindeki metni atlayabilir. ( Şekil 3.8.2) İlgili html belgesinin web sitesi URL'sini, etiketini ve sınıfını ve csv çıktısının adını girmeleri gerekir.

## Scrap

Tag:  Class Name:

Şekil 3.8.2 -Web uygulamasının hurda kısmı

oluşturma sayfasında kullanıcılar sinir ağı modelini otomatik olarak oluşturabilir ve indirebilir. ( Şekil 3.8.3 )

# Create Model

Name

Epoch

Batch Size

Şekil 3.8.3 - Model Oluştur sayfası

Sinir ağı kullanımı sayfasında, kullanıcılar bir model dosyası kullanarak yorum csv'sini analiz edebilir. (Şekil 3.8.4)

# Use Neural Network

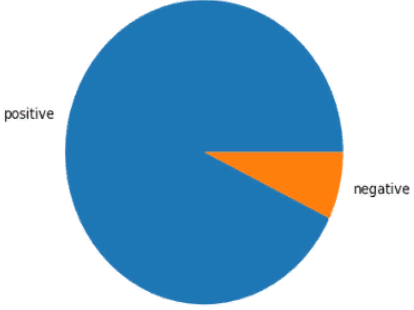
  

Şekil 3.8.4 - Sinir ağı kullan sayfası



Tüm analizörlerde, analiz bittiğinde, kullanıcı ortaya çıkan görüntüyü görebilir ve indirebilir. (Şekil 3.8.5)

### Use Neural Network



positive negative

Download

Browse

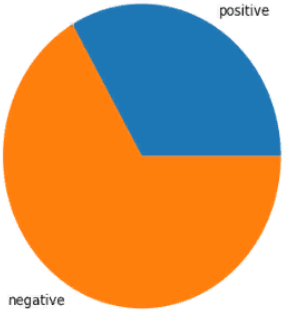
Browse

Plot Name

Analyze

Şekil 3.8.5 -Sayfada yapay sinir ağı kullanılarak yapılan yorumların analizi tamamlanmıştır.

### Use Naive Bayes



positive negative

Download

Browse

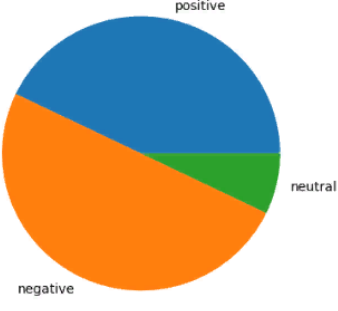
Browse

Plot Name

Analyze

Şekil 3.8.6 -Sayfa Naïve Bayes kullanarak yorumları analiz ederken bitti.

### Use Dictionary



positive negative neutral

Download

Browse

Plot Name

Analyze

Şekil 3.8.7 -Sözlük yöntemi kullanılarak yapılan yorumların analizi bittiğinde sayfa tamamlanır.

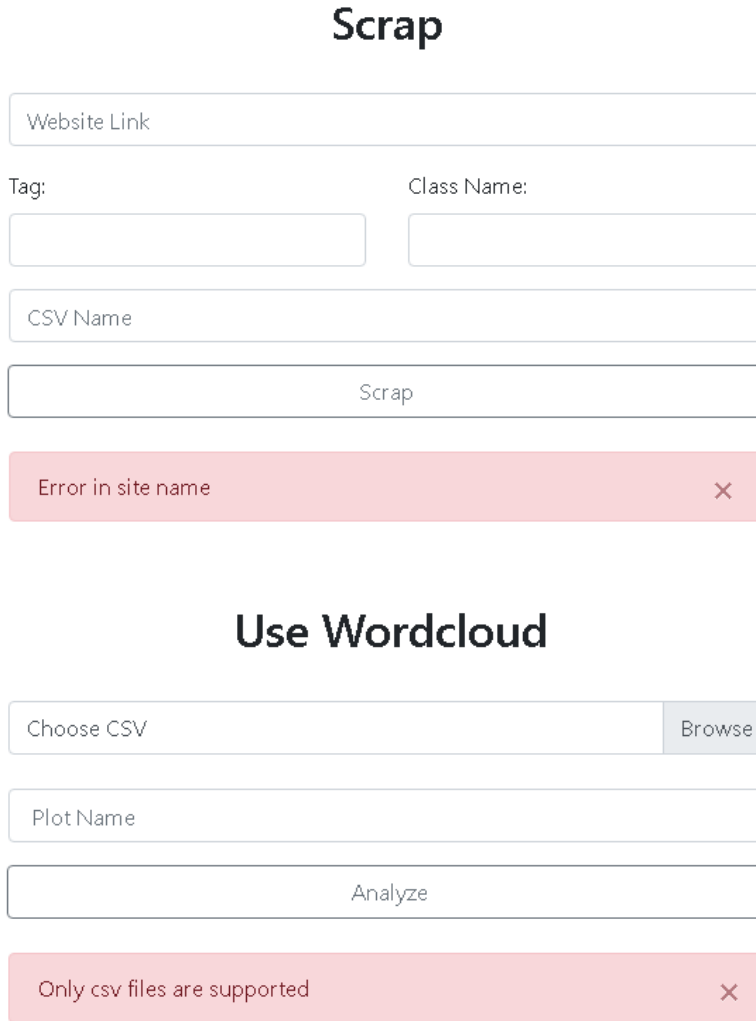
Kullanıcılar ayrıca yorumların kelime bulutunu da görebilirler. (Şekil 3.8.8)



The image shows a word cloud on the left and a web interface on the right. The word cloud features words like 'hate', 'bad', 'awesome', 'good product', 'useless', 'thanks', 'liked', 'heck', 'loved', 'well', and 'awesome'. The web interface is titled 'Use Wordcloud' and includes a 'Browse' button, a 'Plot Name' input field, and an 'Analyze' button. Below the word cloud is a green 'Download' button.

Şekil 3.8.8 -Kelime bulutu kullanılarak yorumların analiz edildiği sayfa bittiğinde

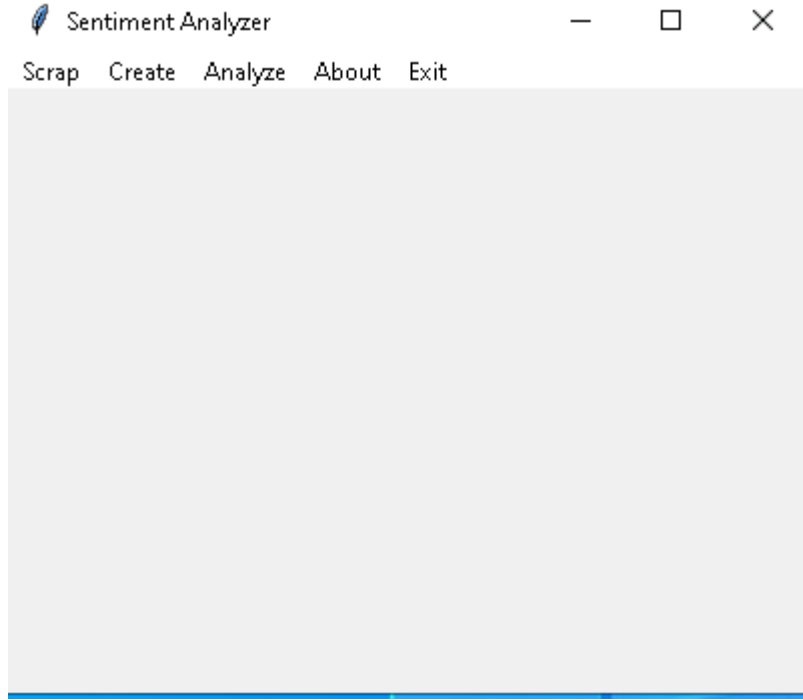
Bir şey eksik veya yanlış olduğunda uygulama ayrıca hata gösterir. (Şekil 3.8.9)



The image shows two error pages. The first is titled 'Scrap' and contains a 'Website Link' input field, 'Tag:' and 'Class Name:' labels with corresponding input fields, a 'CSV Name' input field, and a 'Scrap' button. Below this is a red error message: 'Error in site name' with a close button. The second error page is titled 'Use Wordcloud' and contains a 'Choose CSV' input field with a 'Browse' button, a 'Plot Name' input field, and an 'Analyze' button. Below this is a red error message: 'Only csv files are supported' with a close button.

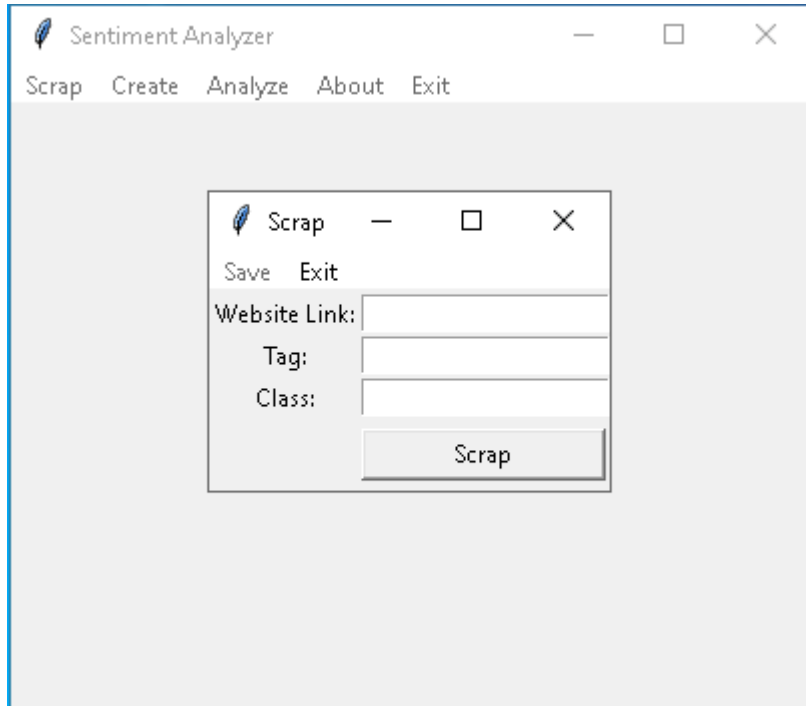
Şekil 3.8.9- Web uygulamasının bazı hata sayfaları

Masaüstü uygulamasının ana sayfası aşağıda gösterilmiştir. Menüde Scrap, Create, Analyze, Hakkında ve Çıkış bölümleri. (Şekil 3.8.10)



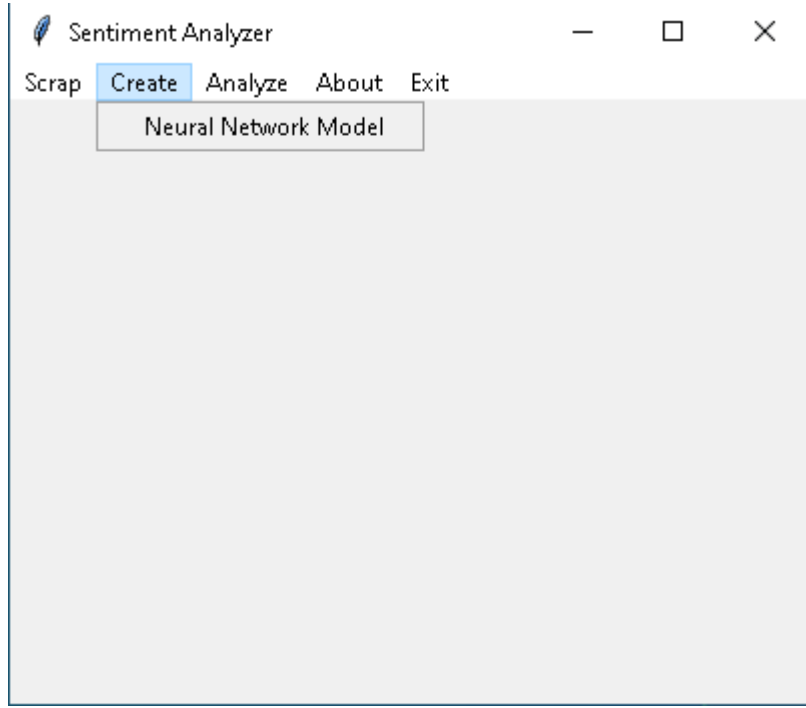
Şekil 3.8.10- Uygulamanın ana penceresi

Hurda sayfasında, kullanıcılar web sitesindeki yorumları kazıyabilir. Kaydet seçeneği aktif değil. Hurda bittiğinde, kullanıcı csv'yi kendi yerel sürücüsüne kaydedebilir. (Şekil 3.8.11)



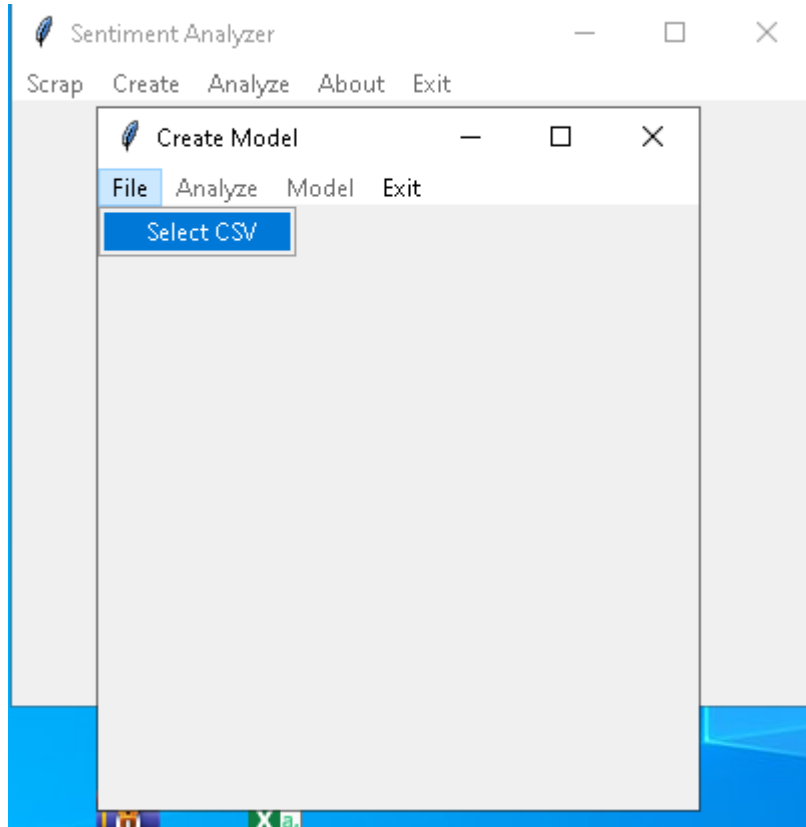
Şekil 3.8.11- Uygulamanın hurda penceresi

penceresinde , kullanıcı bir sinir ağı modeli oluşturabilir. (Şekil 3.8.12)



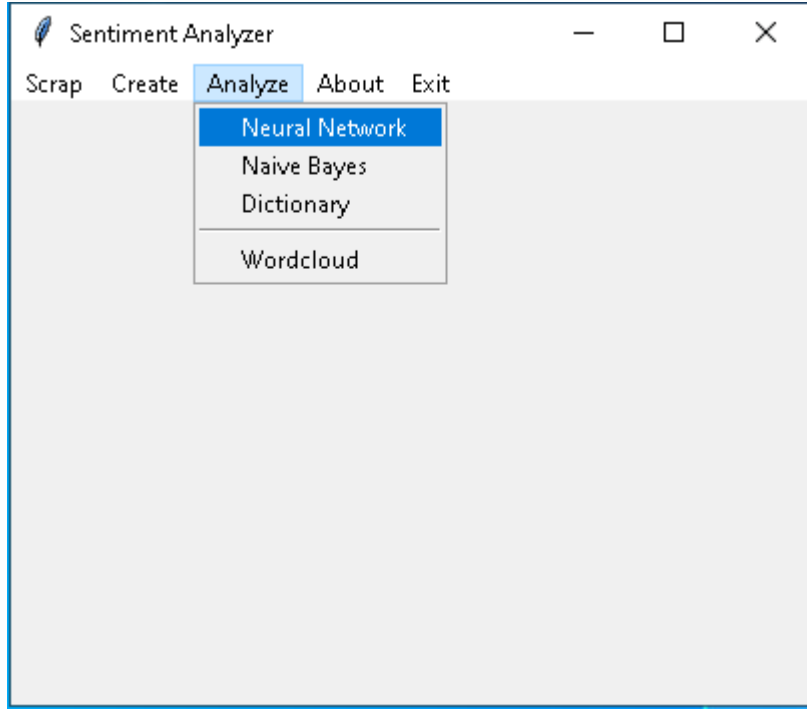
Şekil 3.8.12- Uygulamanın Oluştur seçeneği

Yeni pencerede, kullanıcı bir csv seçer ve her şey yolundaysa, analiz seçeneği erişilebilir hale gelir. Analize tıkladığında ve analiz işlemi bittiğinde, kullanıcılar modele tıklayıp kaydedebilirler. (Şekil 3.8.13)



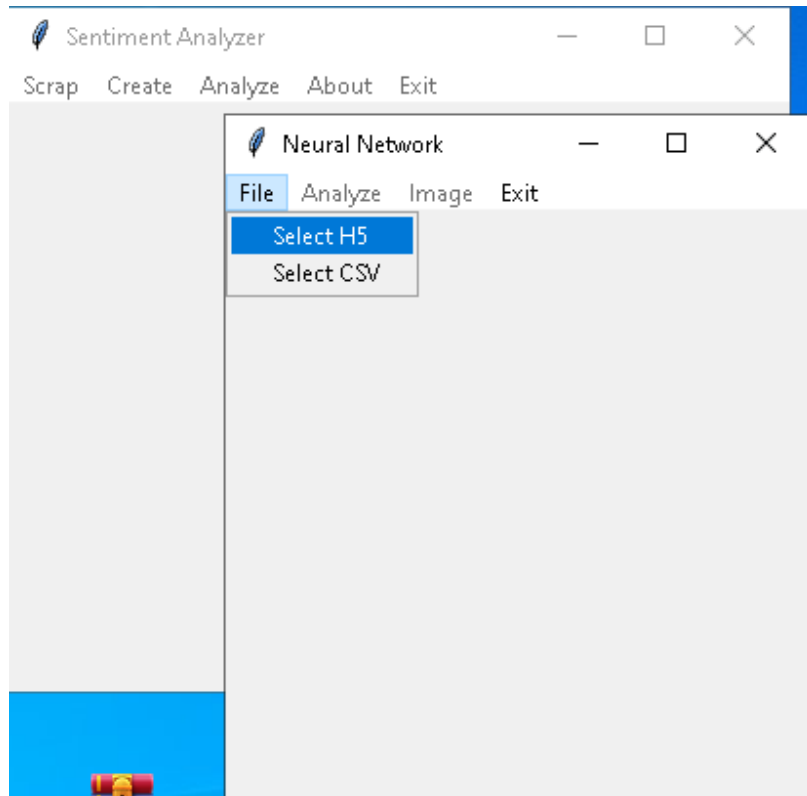
Şekil 3.8.13- Uygulamanın sinir ağı modeli oluştur penceresi

pencerede , kullanıcı bir csv seçer ve her şey yolundaysa, analiz seçeneği erişilebilir hale gelir. Analize tıklandığında ve analiz işlemi bittiğinde, kullanıcılar modele tıklayıp kaydedebilirler. (Şekil 3.8.14)



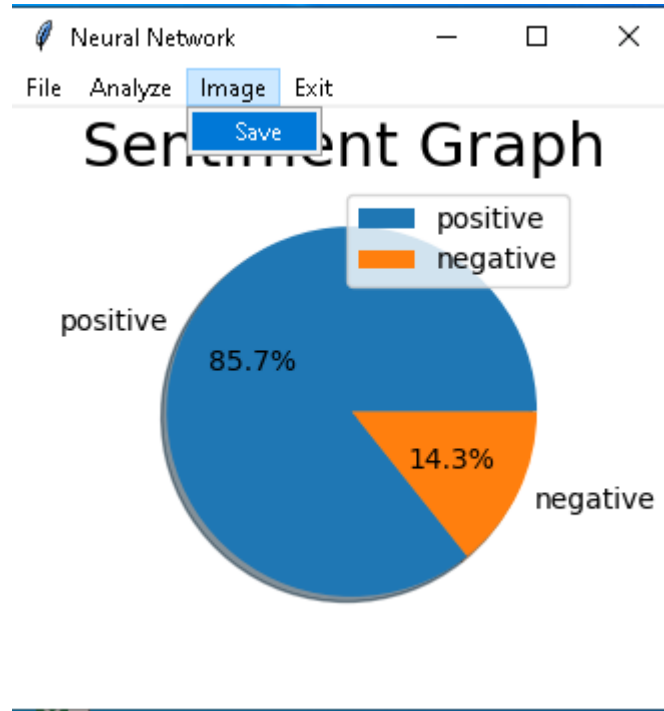
Şekil 3.8.14- Uygulamanın Analiz Et butonu

Sinir ağı bölümünde, kullanıcı h5 dosyasını ve yorumların bulunduğu bir csv dosyasını seçebilir. Ardından analiz seçeneği gelir. (Şekil 3.8.15)



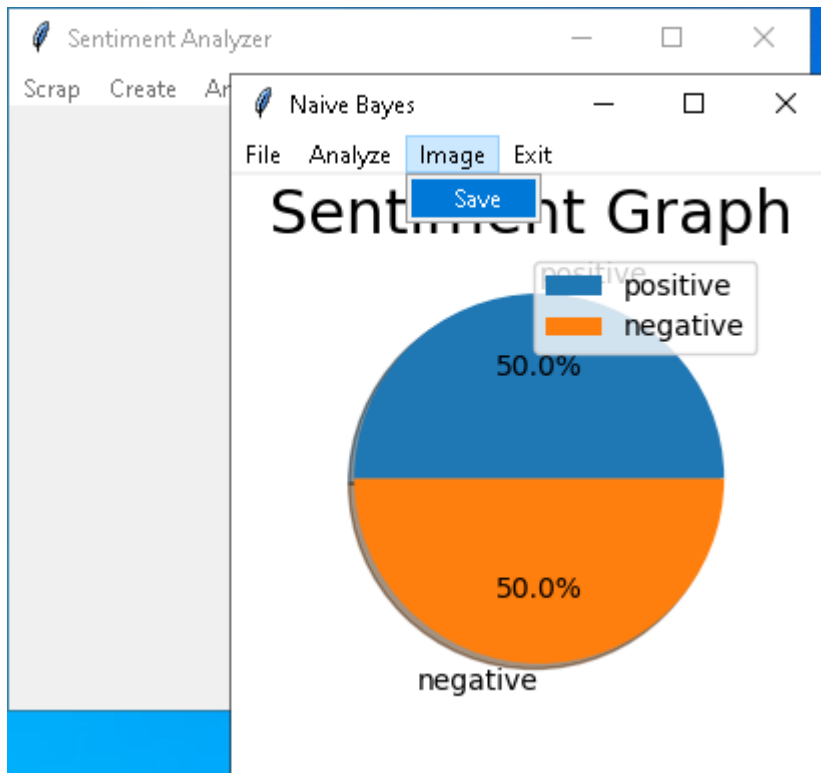
Şekil 3.8.15- Sinir ağı penceresinin Dosya düğmesi

bittiğinde , kullanıcı Görüntü'ye tıklayıp kaydedebilir. (Şekil 3.8.16)



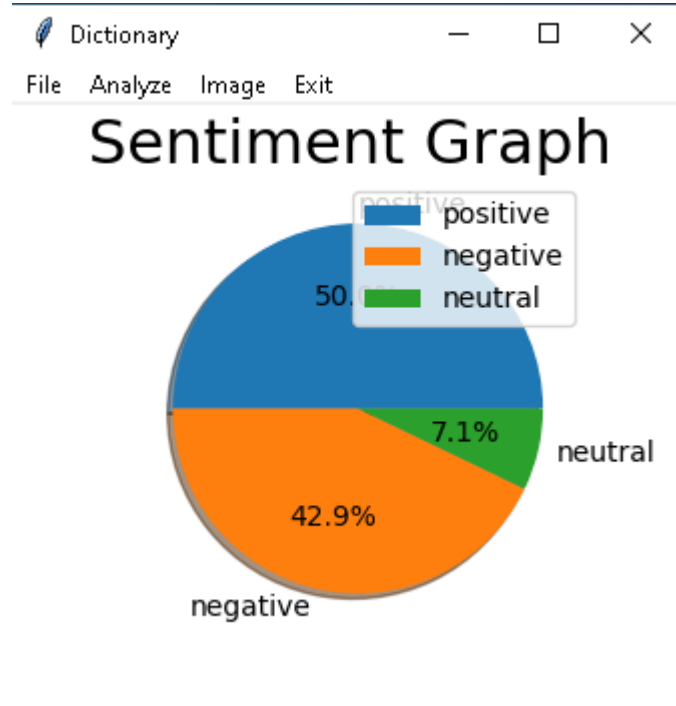
Şekil 3.8.16- Sinir ağı penceresinin görüntüyü kaydet düğmesi

Naïve Bayes kullanarak yorumları analiz ederken, kullanıcının eğitim ve test csv dosyalarını seçmesi gerekir. Sonra gidip Analiz edebilir, görüntüleyebilir ve kaydedebilir. (Şekil 3.8.17)



Şekil 3.8.17- Naïve Bayes penceresinin görüntüyü kaydet düğmesi

Sözlük ve kelime bulutu bölümleri için kullanıcı, yorumların bulunduğu bir csv dosyası seçebilir. Sonra gidip Analiz edebilir, görüntüleyebilir ve kaydedebilir. Ancak sözlük kısmı yalnızca İngilizce içerikleri desteklemektedir. (Şekil 3.8.18), (Şekil 3.8.19)

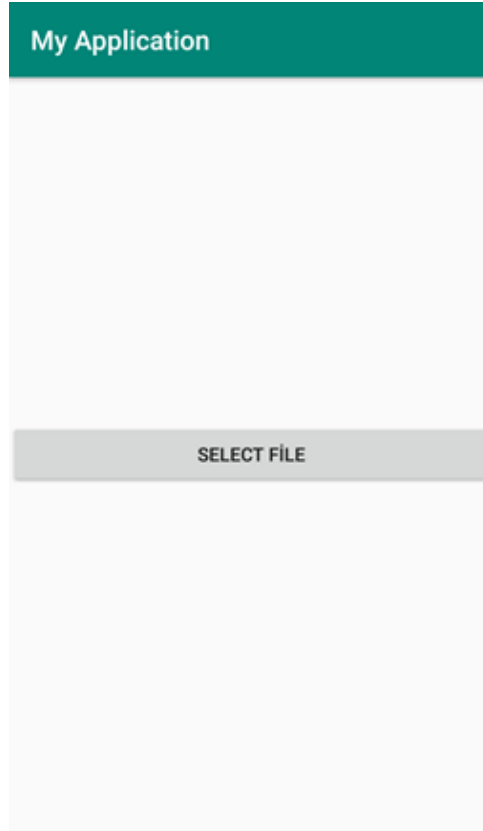


Şekil 3.8.18- Sözlük yöntemi kullanılarak yorumun çözümlenmesi tamamlandığında açılan pencere

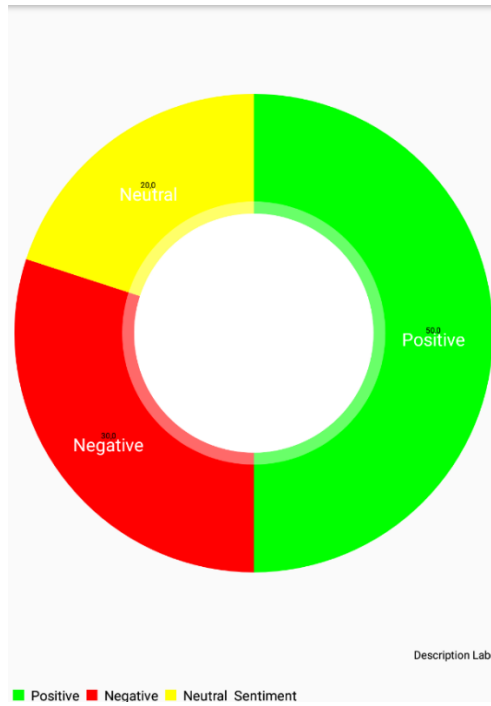


Şekil 3.8.19- Kelime bulutu yöntemi kullanılarak yorumun analiz edilmesinin tamamlandığı pencere

Kullanıcı , mobil uygulamada da grafikleri görebilecek. Grafikler aşağıdaki gibi görünecektir. Örnek pasta grafiği, csv yorumlarının pozitif, negatif ve nötr oranını gösterir. (Şekil 3.8.20)



Şekil 3.8.20- Kullanıcıların analiz etmek için bir csv seçmesi gerekir.



Şekil 3.8.21- Mobil platformda ortaya çıkan pasta grafik



## 4. Sonuç \_

Sonuç olarak Duygu Analizi oldukça faydalı ve önemlidir. Günümüzde insanlar herhangi bir konu hakkındaki düşünce ve fikirlerini topluca paylaşabilmektedirler. Ve web, bu tür bilgilerin ana kaynağıdır. İnternetteki paylaşımlardan yola çıkarak toplumun bir konu hakkında ne düşündüğü belirlenebilir.

Bu projede scraper ve analizör yapılmıştır. Bir kazıyıcı, gönderiyi bir web sitesinde kolayca toplayabilir ve saklayabilir. Öte yandan, toplanan bu verileri analiz etmek için birçok türde analizör eklenmiştir. Ayrıca projeye bir Neural Network model yaratıcısı da ekleniyor. Böylece kullanıcı bu şekilde bir sinir ağı modeli oluşturup sonrasında kullanabilir.

Analizörler farklı yaklaşımlardan seçilir. Bilindiği gibi, duygu analizi ikiye ayrılır. Biri Makine Öğrenimi yaklaşımı, diğeri ise Sözlük Tabanlı yaklaşım. Makine Öğrenimi Yaklaşımının Denetimli ve denetimsiz gibi alt dalları vardır. Denetimli yaklaşımın da alt dalları vardır. Sinir ağı yaklaşımı Lineer sınıflandırıcıya aittir. Ve Naïve Bayes, olasılıksal sınıflandırıcıya aittir. Sözlük tabanlı yaklaşım ise Sözlük Tabanlı yaklaşımın bir alt dalıdır. Bunlara ek olarak, kullanıcı bir kelime bulutu da çizebilir. Kelime bulutu, bir metinde en çok kullanılan kelimeleri görmenin kullanışlı ve kolay yoludur.

Web uygulaması ve masaüstü bölümü için, Hem Windows hem de Linux işletim sistemleri desteklenmektedir. Web uygulaması Linux ve windows server çalıştırabilir. Masaüstü platformu da her iki platformda da kullanılabilir. Mobil uygulama Android cihazlar için geliştirilmiştir.

Diğer sistemlerin aksine, bu programlar herhangi bir web sitesindeki herhangi bir metni kazıyabilir. Ayrıca, sözlük yöntemi dışında, bu program çok dillidir ve herhangi bir dildeki metinleri analiz edebilir. Örneğin, kullanıcı modeli Fransızca kelimelerle eğitirse , Fransızca metinleri analiz edebilir.

## 5. Referanslar

1. Sosyal Ağ API'si, GeeksforGeeks . <https://www.geeksforgeeks.org/social-network-api/>.
2. Web scraping nedir, Scrapinghub : <https://scrapinghub.com/what-is-web-scraping>
3. Machine- Learning yaklaşımı, Kdnuggets : <https://www.kdnuggets.com/2018/03/5-things-sentiment-analysis-classification.html>
4. Sözlüğe Dayalı Yaklaşım, Kdnuggets : <https://www.kdnuggets.com/2018/03/5-things-sentiment-analysis-classification.html>
5. BeautifulSoup , Scrapinghub : <https://info.scrapinghub.com/web-scraping-guide/python-web-scraping-libraries-and-frameworks>
6. Tekrarlayan Sinir Ağları, Wikipedia, <https://www.wikizeroo.org/index.php?q=aHR0cHM6Ly9lbi53aWtpcGVkaWEub3JnL3dpa2kvUmVjdXJyZW50X25ldXJhbF9uZXR3b3Jr>
7. RNN'lere giriş, wildml , <http://www.wildml.com/2015/09/recurrent-neural-networks-tutorial-part-1-introduction-to-rnns/>
8. Naïve Bayes, Orta, <https://medium.com/@martinpella/naive-bayes-for-sentiment-analysis-49b37db18bf8>
9. Şişe, Github , <https://github.com/pallets/flask>
10. Tkinter , Python, <https://docs.python.org/2/library/tkinter.html>
11. Bulut , Faruk. "Kuantize edilmiş Haar dalgacık dönüşümü yoluyla düşük dinamik aralıklı histogram eşitleme (LDR-HE)." *Görsel Bilgisayar* 38, no. 6 (2022): 2239-2255.

## 6. Terimler Listesi

### 6.1. Tanımlar

*Çerçeve* - uygulama geliştirmek için bir platformdur. Geliştiricilerin belirli bir platform için yazılım geliştirebilecekleri bir ortam sağlar.

*Sözlük* - belirli bir lehçede veya konuda kullanılan tüm kelimeler veya bir kelime referansı.

### 6.2. Kısaltmalar

<b>API</b>	Uygulama programlama Arayüzü
<b>GUI</b>	Grafiksel kullanıcı arayüzü
<b>HTML</b>	Köprü Metni İşaretleme Dili
<b>xml</b>	Genişletilebilir İşaretleme Dili
<b>LSTM</b>	Uzun kısa süreli hafıza
<b>kullanıcı arayüzü</b>	Kullanıcı arayüzü