

E-Ticaret Müşteri Yorumları İle Duygu Analizi

miuul

PROJE SUNUM İÇERİĞİ

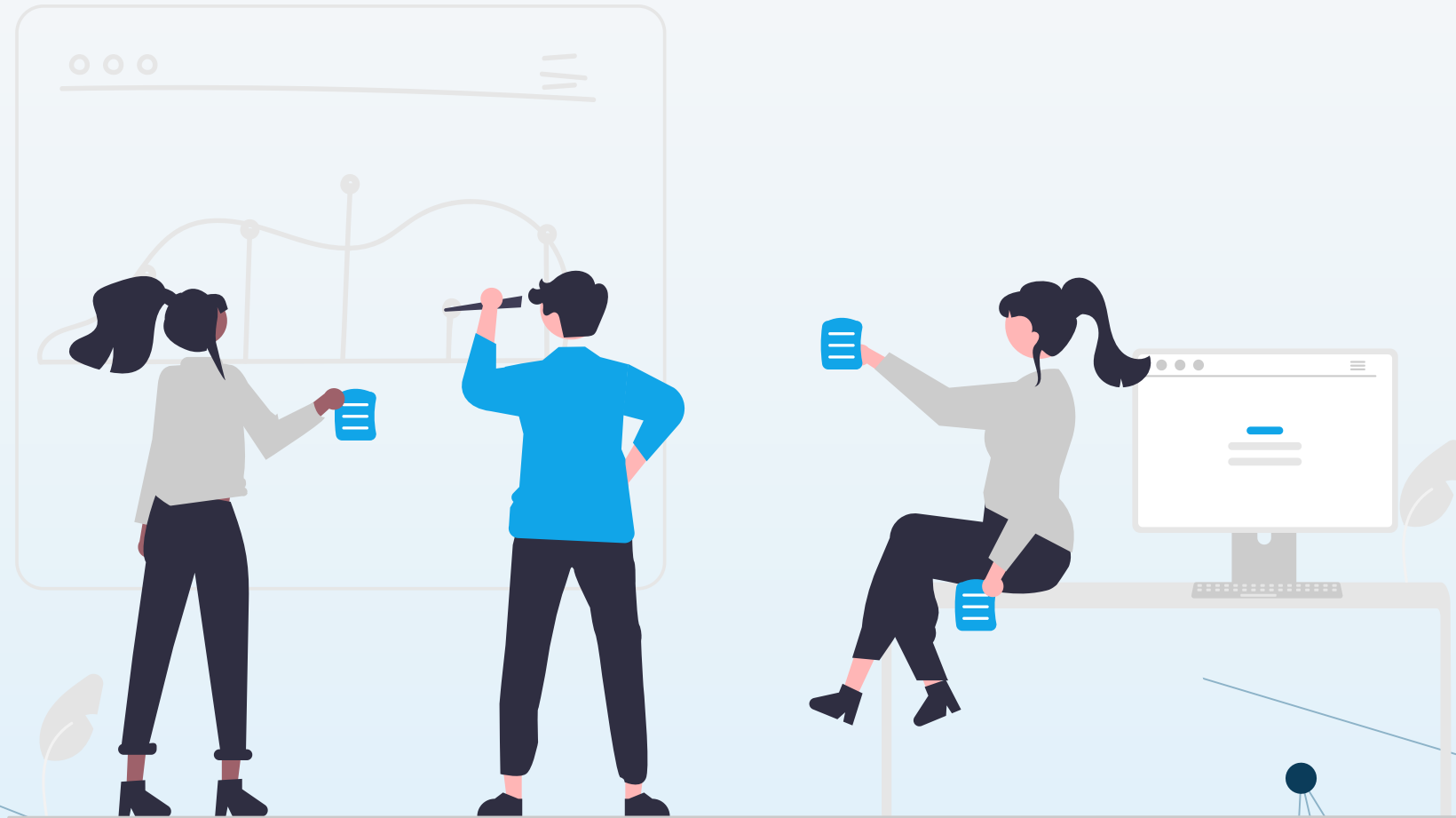
01. İş İhtiyacı

02. Proje Amacı

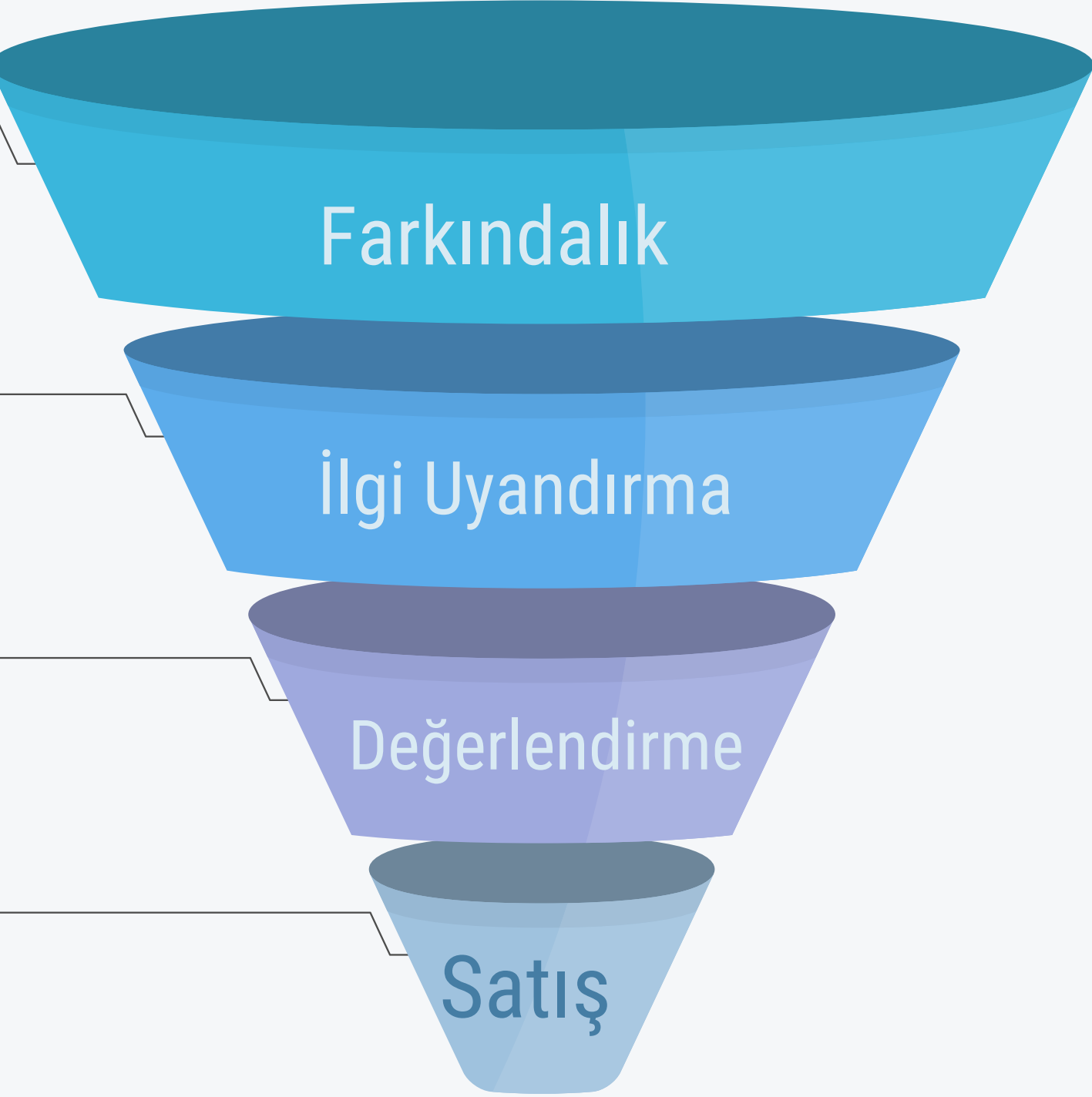
03. Proje Zaman Planlaması

04. Proje Tanıtımı

05. Proje Ekibi

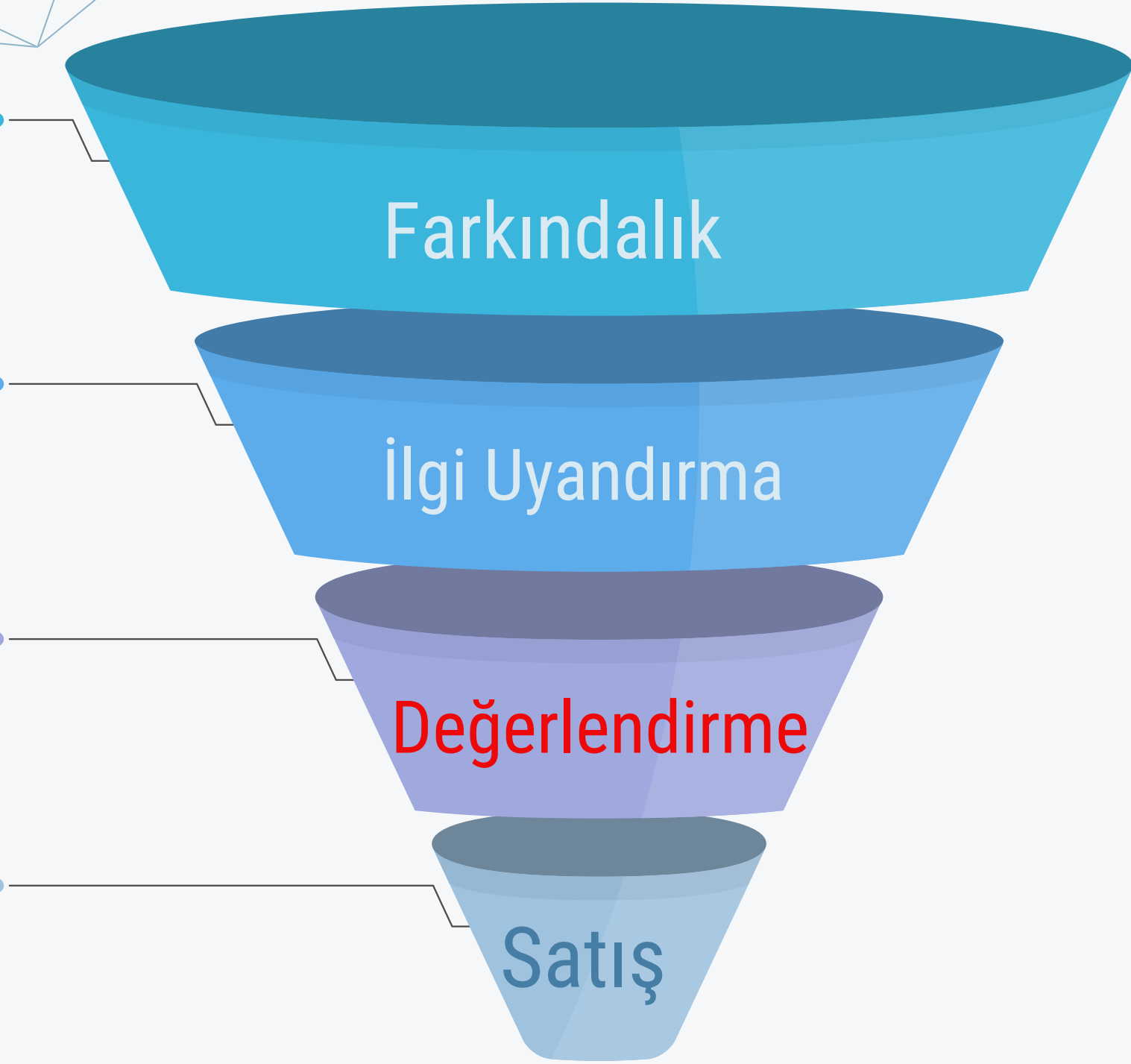


iř İHTİYACI



- Müşteri bir ürünü almaya karar vermeden önce içsel bir yolculuktan geçmektedir. Bu durum pazarlama dünyasında satış hunisi (sales funnel) olarak tanımlanmaktadır.
- Müşterinin karar verme sürecini en çok etkileyen ve zaman alan süreç değerlendirme süreci olmaktadır.

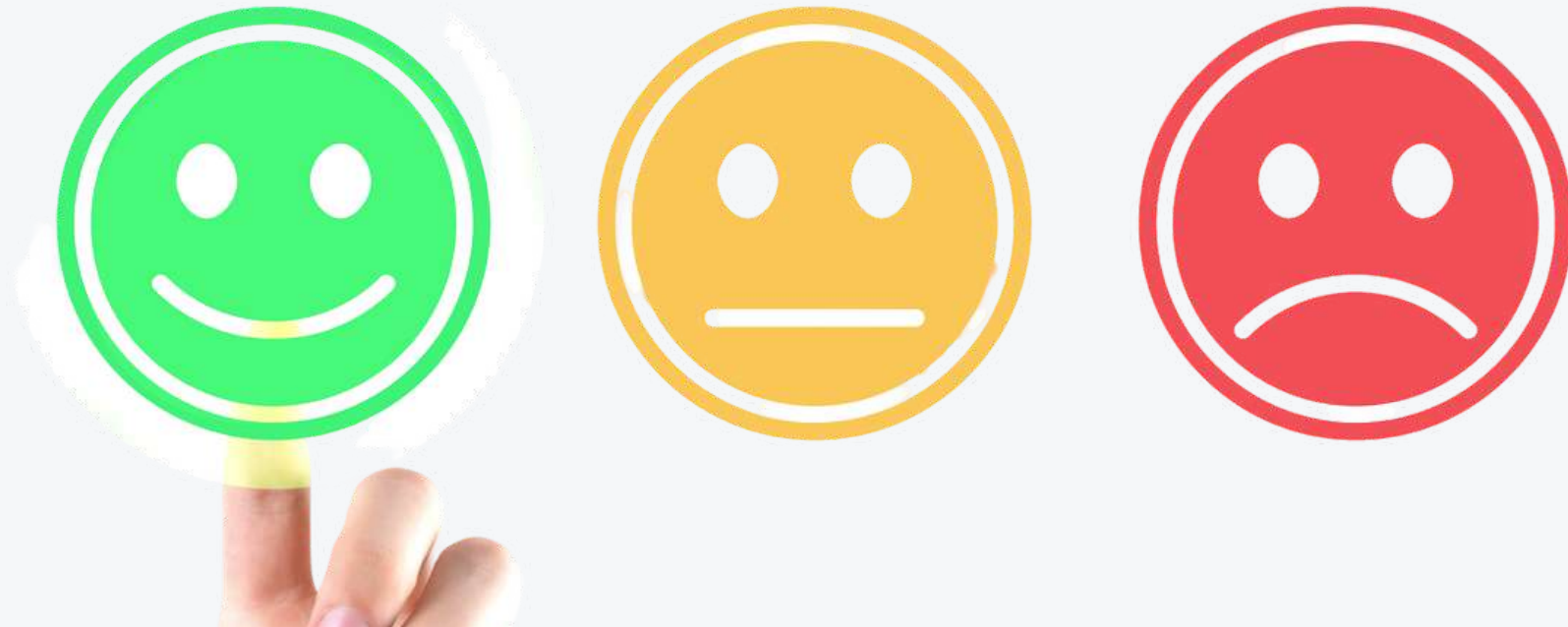
İŞ İHTİYACI



- Değerlendirme sürecinde müşterinin daha hızlı ve etkin karar almasının sağlamak için ürünlerin iyi ve kötü yönlerinin hızlı şekilde karşılaştırmasını yapabilecek bir yazılım ihtiyacı doğmuştur.

PROJENİN AMACI

- Projenin amacı; satış hunisinde değerlendirme sürecinde olan müşterinin ürün bilgilerini uygulamamıza girerek bu ürünün pazar yerlerinde ne kadar olumlu olumsuz yorum aldığını, bu yorumlarda iyi ve kötü yönlerinin neler olduğunu hızlı şekilde tespit ederek, doğru ve hızlı karar almasını sağlamaktır.



PROJE TANITIMI

- Müşteriye ürün yorumlarından anlamlı bir sonuç çıkarmak için öncelikle metinden duygu analizi (Sentiment Analysis) yapmayı hedefledik.
- Bu proje için farklı pazar yerlerinden akıllı robot süpürge kategorisinde en popüler dört markaya ait farklı modeller belirleyerek, bu modellere ait 12,000 adet yorumu veri olarak aldık.

PROJE TANITIMI

Veri Seti

4

MARKA

10

MODEL

12.000

YORUM



PROJE TANITIMI

Veri Seti

Değişkenler

Marka : Ürün markası.

Model : Ürün modeli.

Fiyat : Ürünün güncel fiyatı.

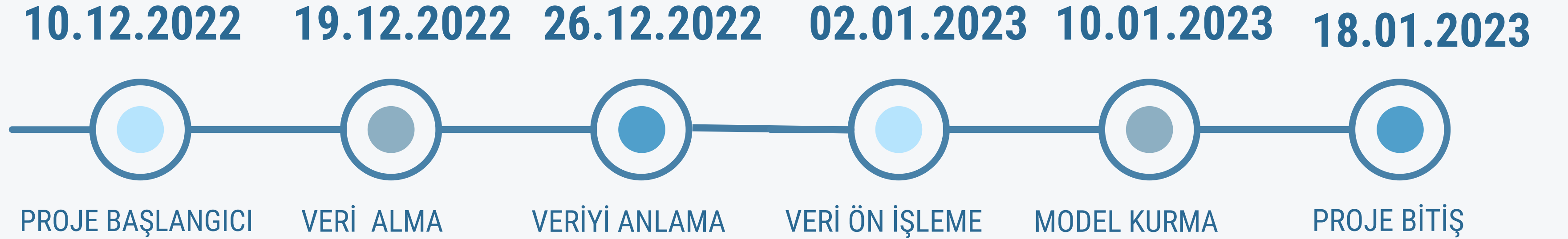
Puan : Müşteri tarafından ürüne verilen puan.

Yorum Tarihi : Müşteri tarafından yorum yapılma tarihi.

Yorum : Müşteri tarafından yapılan yorum.



Proje Zaman Planlaması



PROJE TANITIMI

Veriyi Anlama



- Bu aşamada veriyi inceleyerek değişkenlerin neler olduğunu değerlendirdik.
- Özellikle projede hedef değişken olarak kullanacağımız müşteri yorumlarını inceleyerek yorumların derecelendirme (rate) ile tutarlı olup olmadığını ve metinden temizlenmesi gereken alanları belirledik.

PROJE TANITIMI

Veri Ön İşleme

- Veri ön işleme aşamasında; öncelikle veride hangi bilgilerin olması gerektiğini karar verdik.
- Veriyi belli standartlar belirleyerek bu standartlara göre tekrar biçimlendirdik.
- Metin analizinde fazladan değişken oluşmasını engellemek ve doğruluğu artırmak için verideki noktalama işaretleri, emoji ve herhangi bir anlam ifade etmeyen kelimeleri veriden çıkardık.

PROJE TANITIMI

Veri Ön İşleme

Kullanılan Kütüphaneler



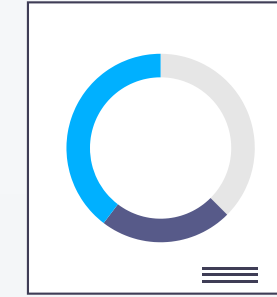
matplotlib



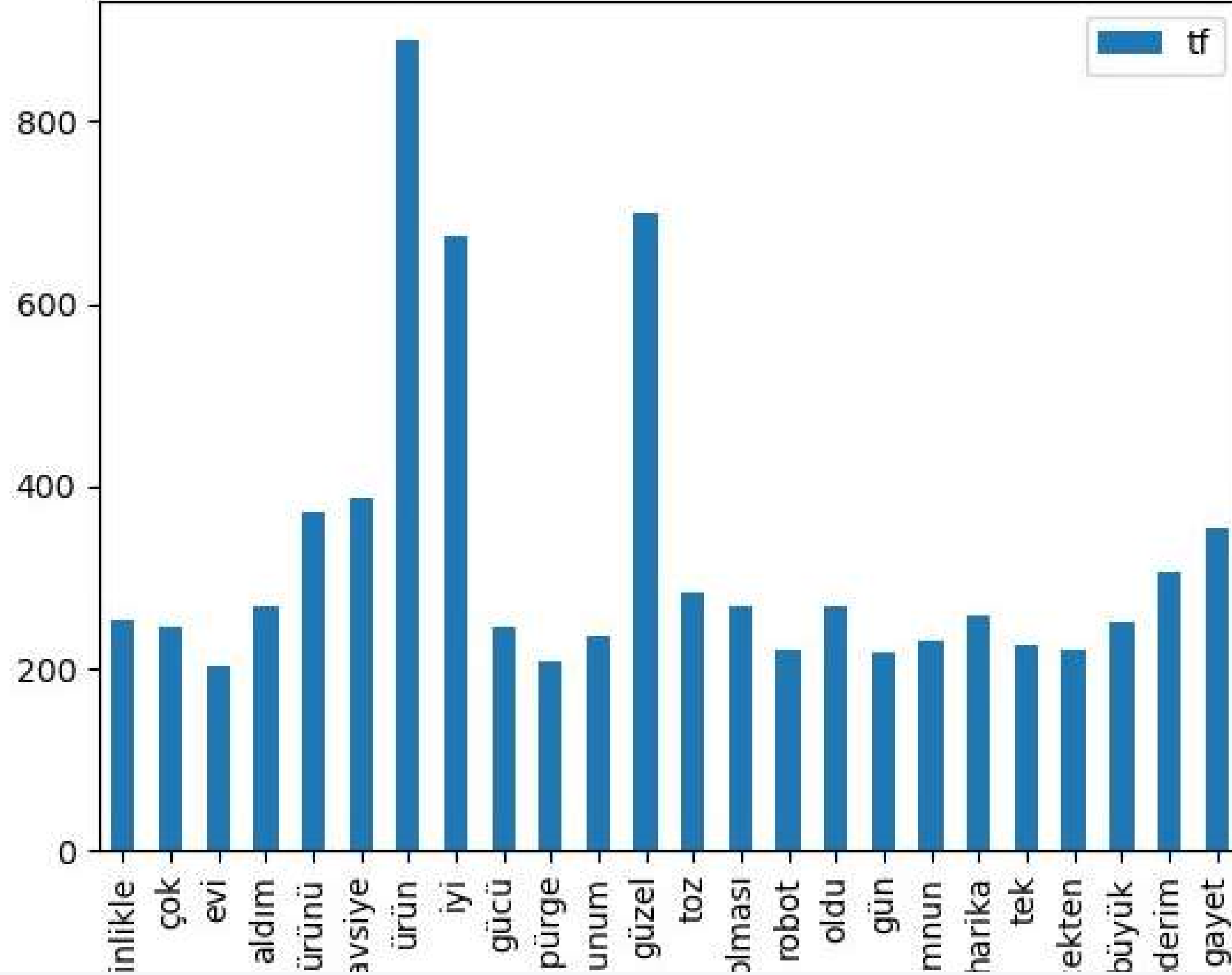
PROJE TANITIMI

Veri Ön İşleme

- Müşteri yorumlarında geçen kelimelerin frekanslarını bir comprehension yapısı kullanarak belirledik ve en çok geçen kelimeleri tespit ederek bar plot ve wordcloud ile görselleştirdik.



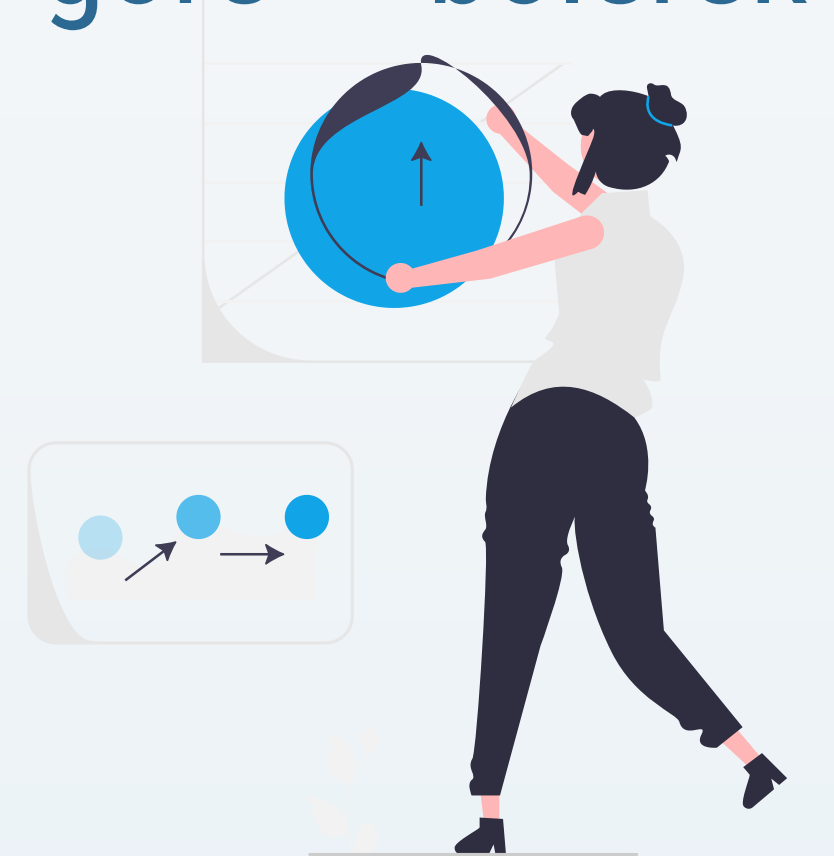
Kelime Frekansları



PROJE TANITIMI

Model Kurma

- Makine öğrenmesinde dengesiz veri dağılımını düzenlemek ve farklı veri ile modeli deneyebilmek için bir comprehension yapısı kullanarak müşteri yorumlarını rate'lere göre bölerek ağırlıklandırdık.





PROJE TANITIMI

Model Kurma

- Müşteri yorumlarını makine öğrenmesi modeline sokabilmek için Sentiment Analyzer ile yorumların pozitif ve negatif sınıfa ait olma olasılıklarını hesapladık.
- Hesaplanan olasılıkları Sentiment Label ile pozitif ve negatif olarak sınıflandırarak, bu sınıfları Label Encoder ile pozitif değerlere 1 negatif değerlere 0 olarak atadık.

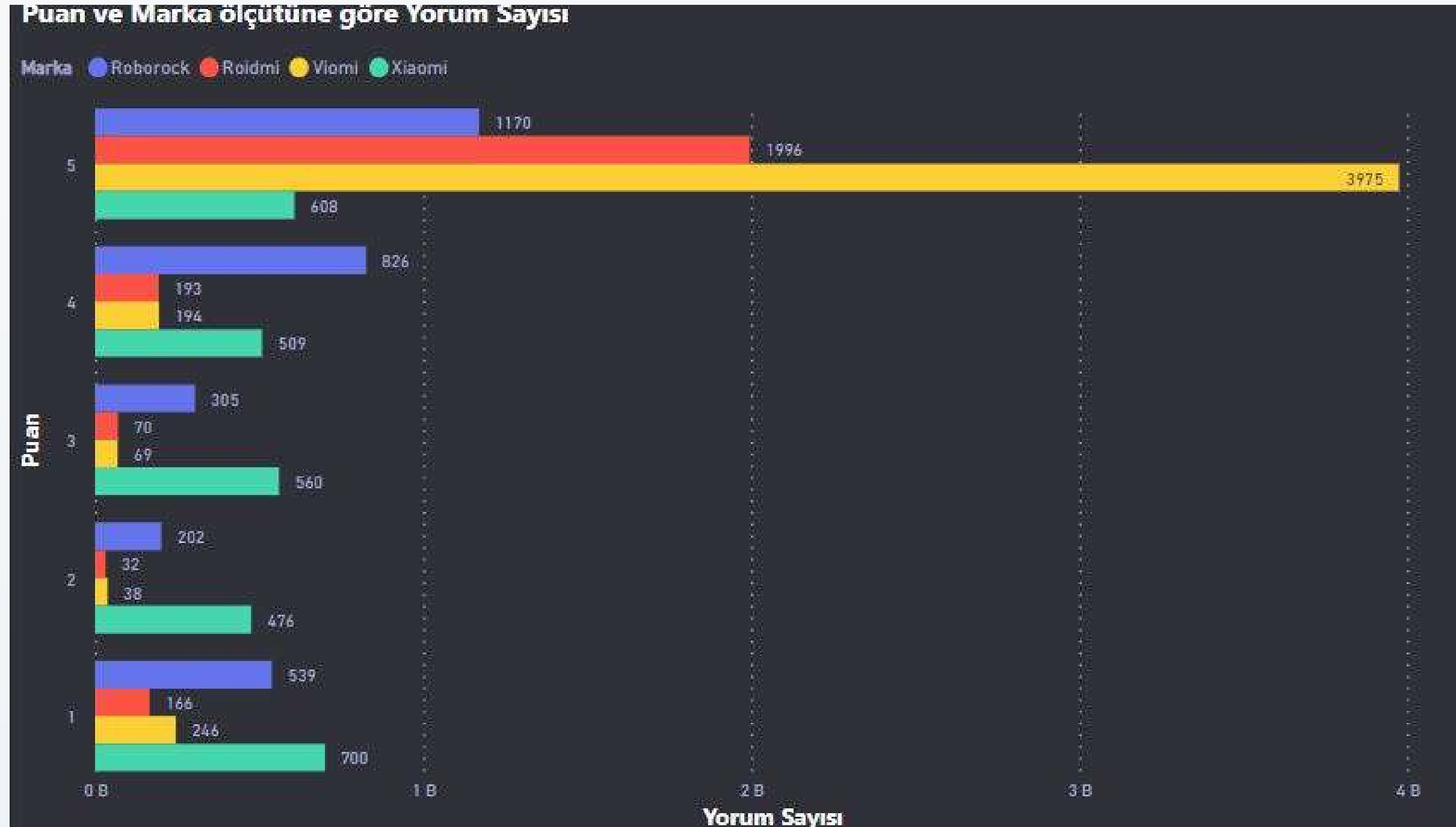
PROJE TANITIMI

Marka ve Modele Göre Pozitif-Negatif Oranları



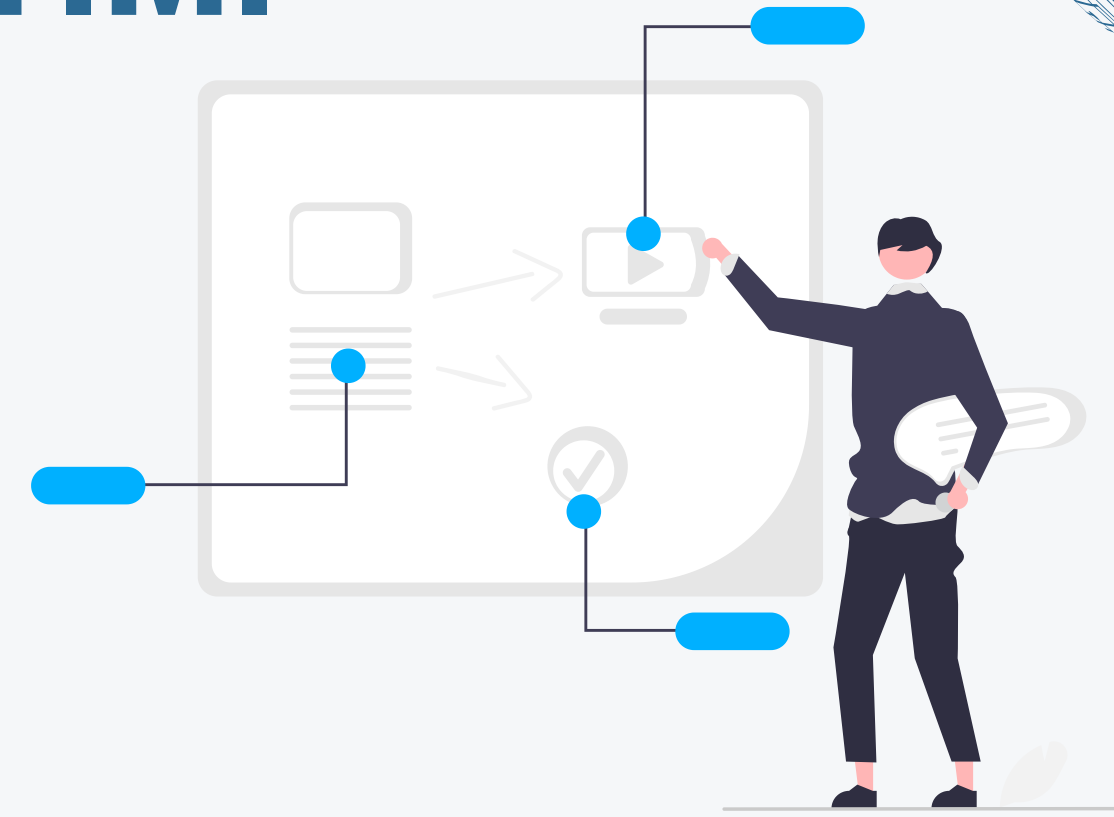
PROJE TANITIMI

Marka ve Puana Göre Yorum Dağılımları



PROJE TANITIMI

Model Kurma



- Müşteri yorumlarında geçen her kelimeyi Tf-Idf yöntemi standartlaştırdık.
- Oluşturulan değişkenleri RandomForest, Logistic Regression ve Light GBM makine öğrenimi modellerinde test ettik ve sonraki yansıda belirtilen doğruluk oranlarını elde ettik.

PROJE TANITIMI

BASE MODEL

Model

RandomForest

Logistic Regression

Light GBM

Accuracy

0,85

0,83

0,82



PROJE TANITIMI

Model Kurma

- Test ettiğimiz modellerden doğruluk oranı en yüksek olan model ile veriyi tekrar test ederek bu model üzerinde hiper parametre optimizasyonu yaparak model doğruluğunu artırdık.
- Müşteri yorumlarının duygu analizini RandomForest makine öğrenimi modeli ile 0,92 oranında doğruluk ile tahmin ederek projemizi başarı ile sonuçlandırdık.

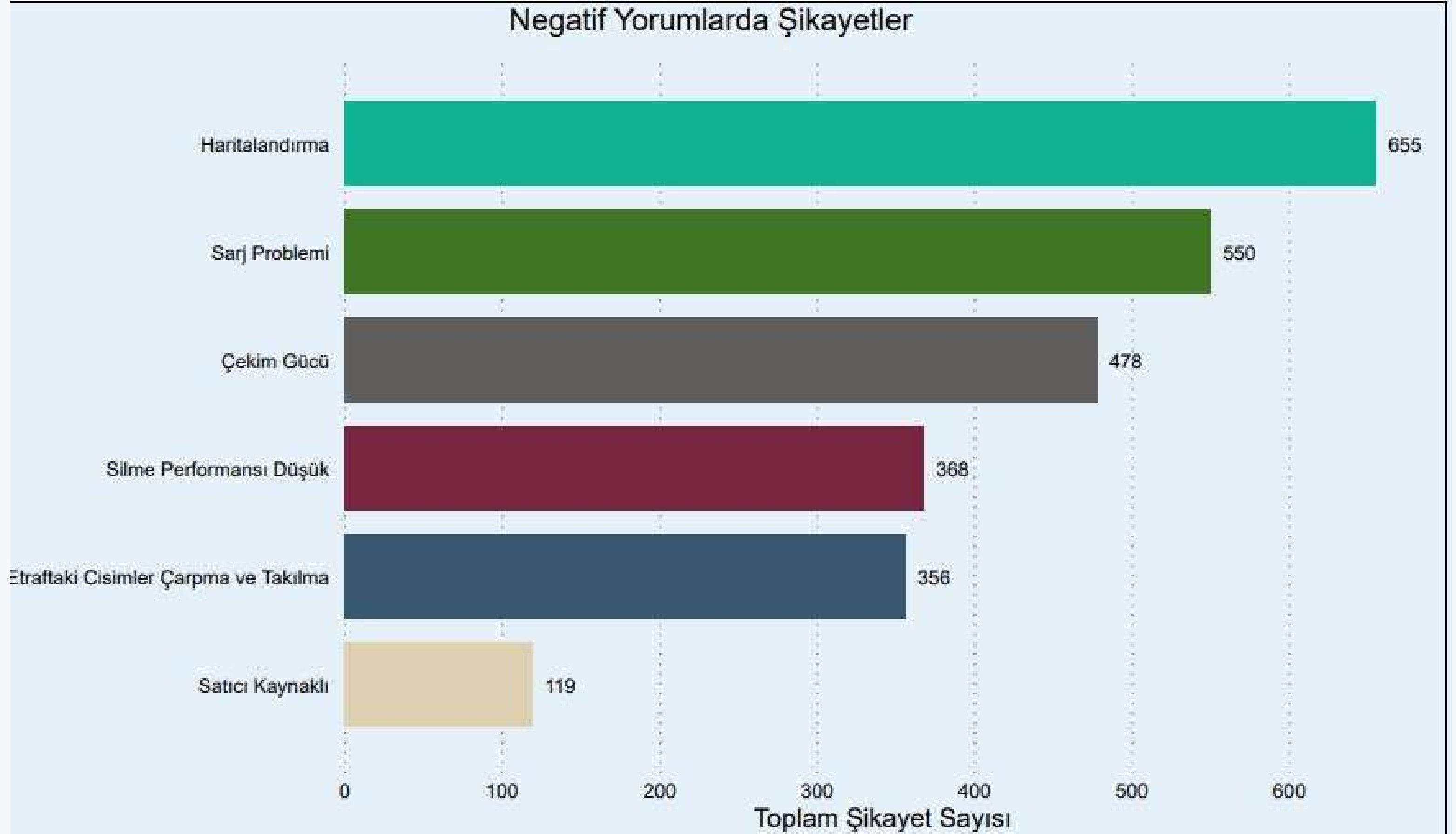
PROJE TANITIMI



Power BI Görselleştirme



PROJE TANITIMI





Altuğ Gürgöl



Furkan Aslan



Nur Demircioğlu



Fatih Songur

PROJE EKİBİ

**DİNLEDİĞİNİZ İÇİN
TEŞEKKÜRLER**



Proje Web Uygulaması

Projemizin kod altyapısını tamamladıktan sonra nihai son kullanıcıya ulaştırmak için streamlit'de bir arayüz tasarlayarak müşterilerin kullanımına sunduk.

