



Home

Service

About Us

Contact

YAPAY ZEKA MÜHENDİSLİĞİ

Vaka Çalışma Sunumu

İstanbul Teknik Üniversitesi
Oğuzhan Atmaca



Next Slide →



İLK ADIM; YAPILACAK TESPİTİ

PROBLEM

Employee attrition oranı, şirketin operasyonel verimliliğini düşürmüştür ve maliyetlerini artırmıştır. Bu nedenle yöneticiler, çalışan kaybının nedenlerini anlayarak önleyici stratejiler geliştirmek istemektedir.



HEDEF / AMAÇ

Çalışan kaybının nedenlerini veri analiziyle belirleyip, önleyici stratejiler geliştirmek amacıyla bir karar destek sistemi oluşturmak.



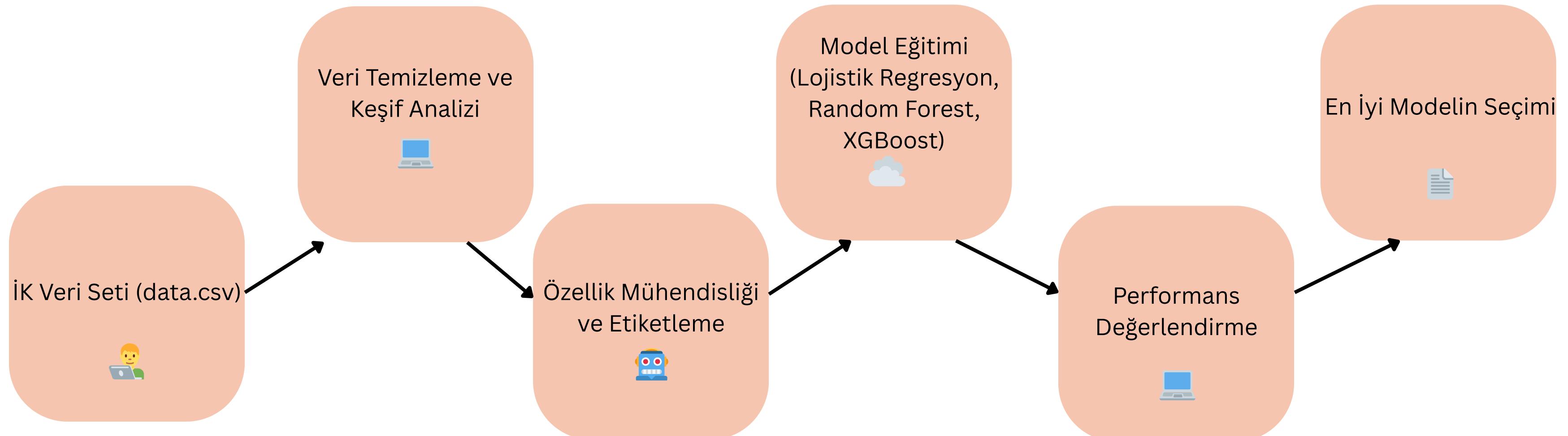
ÇÖZÜM

Veri seti analiz edilerek önemli faktörler belirlenecek, eksik/aykırı veriler tespit edilecek ve çalışan ayrılma ihtimalini tahmin eden sınıflandırma modelleri (örneğin Logistic Regression ve Random Forest) oluşturulacak, modeller Accuracy, Precision, Recall, F1-Score ve ROC-AUC metrikleriyle karşılaştırılacaktır.

	Attrition	Count
1	No	1233
2	Yes	237



İLK VAKA İÇİN ÇÖZÜM PLANI





ÇALIŞAN KAYBINI ETKİLEYECEK TEMEL FAKTÖRLER

1. İş Tatmini:

İşten Ayrılanların Ortalama İş Tatmini: 2.47

Kalanların Ortalama İş Tatmini: 2.78

İstatistiksel Anlamlılık (p-değeri): 0.0001

2. Aylık Maaş:

İşten Ayrılanların Ortalama Maaşı: 4787.09

Kalanların Ortalama Maaşı: 6832.74

İstatistiksel Anlamlılık (p-değeri): 0.0000

3. Şirkette Geçirilen Süre:

İşten Ayrılanların Ortalama Şirket Süresi: 5.13 yıl

Kalanların Ortalama Şirket Süresi: 7.37 yıl

İstatistiksel Anlamlılık (p-değeri): 0.0000

4. Çalışma-Yaşam Dengesi:

İşten Ayrılanların Ortalama Denge Skoru: 2.66

Kalanların Ortalama Denge Skoru: 2.78

İstatistiksel Anlamlılık (p-değeri): 0.0142

5. Departman Bazlı Çalışan Kaybı Oranları:

Kalanlar (%) Ayrılanlar (%)

Department	Kalanlar (%)	Ayrılanlar (%)
Human Resources	80.952381	19.047619
Research & Development	86.160250	13.839750
Sales	79.372197	20.627803

6. İş Rolü Bazlı Çalışan Kaybı Oranları:

Kalanlar (%) Ayrılanlar (%)

JobRole	Kalanlar (%)	Ayrılanlar (%)
Healthcare Representative	93.129771	6.870229
Human Resources	76.923077	23.076923
Laboratory Technician	76.061776	23.938224
Manager	95.098039	4.901961
Manufacturing Director	93.103448	6.896552
Research Director	97.500000	2.500000
Research Scientist	83.904110	16.095890
Sales Executive	82.515337	17.484663
Sales Representative	60.240964	39.759036

ÖNEMLİ BULGULAR VE ÖNERİLER

1. İş Tatmini:

- İş tatmini düşük olan çalışanlarda ayrılma oranı daha yüksektir
- Öneri: Düzenli iş tatmini anketleri ve geri bildirim mekanizmaları kurulmalı

2. Maaş ve Kariyer Gelişimi:

- Maaş artışı ve kariyer gelişimi fırsatları çalışan bağılılığını etkilemektedir
- Öneri: Şeffaf maaş politikaları ve kariyer gelişim planları oluşturulmalı

3. Çalışma-Yaşam Dengesi:

- Denge skoru düşük olan çalışanlarda ayrılma eğilimi daha yüksektir
- Öneri: Esnek çalışma saatleri ve uzaktan çalışma seçenekleri sunulmalı

4. Departman ve İş Rolü:

- Bazı departman ve iş rollerinde ayrılma oranları daha yüksektir
- Öneri: Yüksek ayrılma oranı olan departmanlarda özel retansiyon stratejileri geliştirilmeli

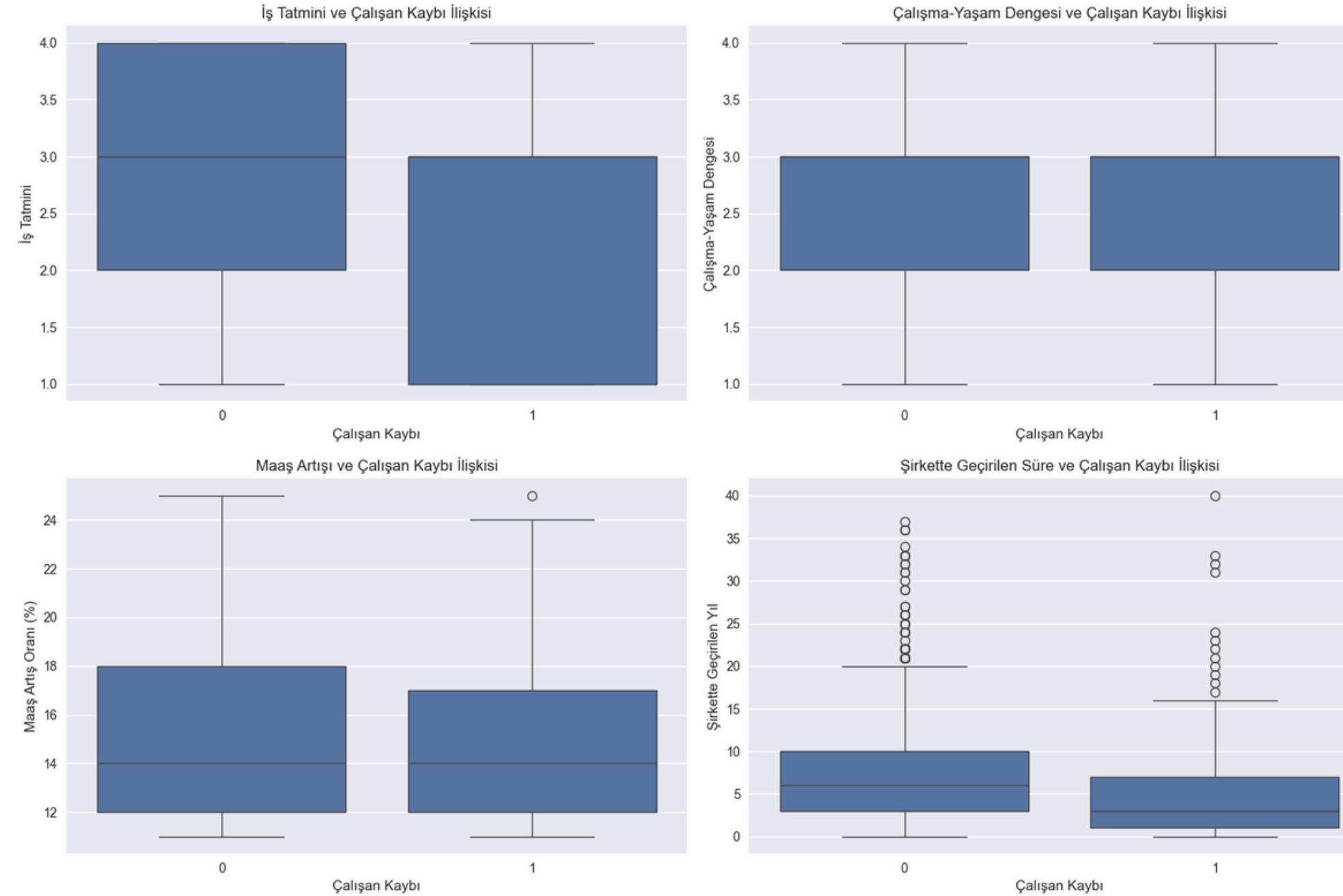
5. Şirkette Geçirilen Süre:

- Yeni çalışanlarda ayrılma oranı daha yüksektir
- Öneri: Yeni çalışan oryantasyon programları güçlendirilmeli

Analizde çalışan kaybını etkileyen faktörler t-testi, korelasyon ve oran analizleriyle incelendi. Sayısal farklar için boxplot, kategorik değişkenler için barplot ve korelasyonlar için heatmap kullanıldı. Elde edilen bulgular doğrultusunda ayrılma riski yüksek gruplar belirlendi ve hedefe yönelik öneriler sunuldu.



ÇALIŞAN KAYBINI ETKİLEYECEK TEMEL FAKTÖRLER



Sonraki →



VERİ ANALİZİ

Aykırı Veri İstatistikleri:

	Alt Sınır	Üst Sınır	Aykırı Veri Sayısı	...	Ortalama	Medyan	Standart Sapma
Age	10.50	62.50	0.0	...	36.92	36.0	9.14
DailyRate	-573.00	2195.00	0.0	...	802.49	802.0	403.51
DistanceFromHome	-16.00	32.00	0.0	...	9.19	7.0	8.11
Education	-1.00	7.00	0.0	...	2.91	3.0	1.02
EmployeeCount	1.00	1.00	0.0	...	1.00	1.0	0.00
EnvironmentSatisfaction	-1.00	7.00	0.0	...	2.72	3.0	1.09
HourlyRate	-5.62	137.38	0.0	...	65.89	66.0	20.33
JobInvolvement	0.50	4.50	0.0	...	2.73	3.0	0.71
JobLevel	-2.00	6.00	0.0	...	2.06	2.0	1.11
JobSatisfaction	-1.00	7.00	0.0	...	2.73	3.0	1.10
MonthlyIncome	-5291.00	16581.00	114.0	...	6502.93	4919.0	4707.96
MonthlyRate	-10574.75	39083.25	0.0	...	14313.10	14235.5	7117.79
NumCompaniesWorked	-3.50	8.50	52.0	...	2.69	2.0	2.50
PercentSalaryHike	3.00	27.00	0.0	...	15.21	14.0	3.66
PerformanceRating	3.00	3.00	226.0	...	3.15	3.0	0.36
RelationshipSatisfaction	-1.00	7.00	0.0	...	2.71	3.0	1.08
StandardHours	80.00	80.00	0.0	...	80.00	80.0	0.00
StockOptionLevel	-1.50	2.50	85.0	...	0.79	1.0	0.85
TotalWorkingYears	-7.50	28.50	63.0	...	11.28	10.0	7.78
TrainingTimesLastYear	0.50	4.50	238.0	...	2.80	3.0	1.29
WorkLifeBalance	0.50	4.50	0.0	...	2.76	3.0	0.71
YearsAtCompany	-6.00	18.00	104.0	...	7.01	5.0	6.13
YearsInCurrentRole	-5.50	14.50	21.0	...	4.23	3.0	3.62
YearsSinceLastPromotion	-4.50	7.50	107.0	...	2.19	1.0	3.22
YearsWithCurrManager	-5.50	14.50	14.0	...	4.12	3.0	3.57

[25 rows x 9 columns]

En Çok Aykırı Veri İçeren Değişkenler (Top 5):

	Aykırı Veri Sayısı	Aykırı Veri Yüzdesi
TrainingTimesLastYear	238.0	16.19
PerformanceRating	226.0	15.37
MonthlyIncome	114.0	7.76
YearsSinceLastPromotion	107.0	7.28
YearsAtCompany	104.0	7.07

Veri setinde herhangi bir eksik veri tespit edilmemiştir

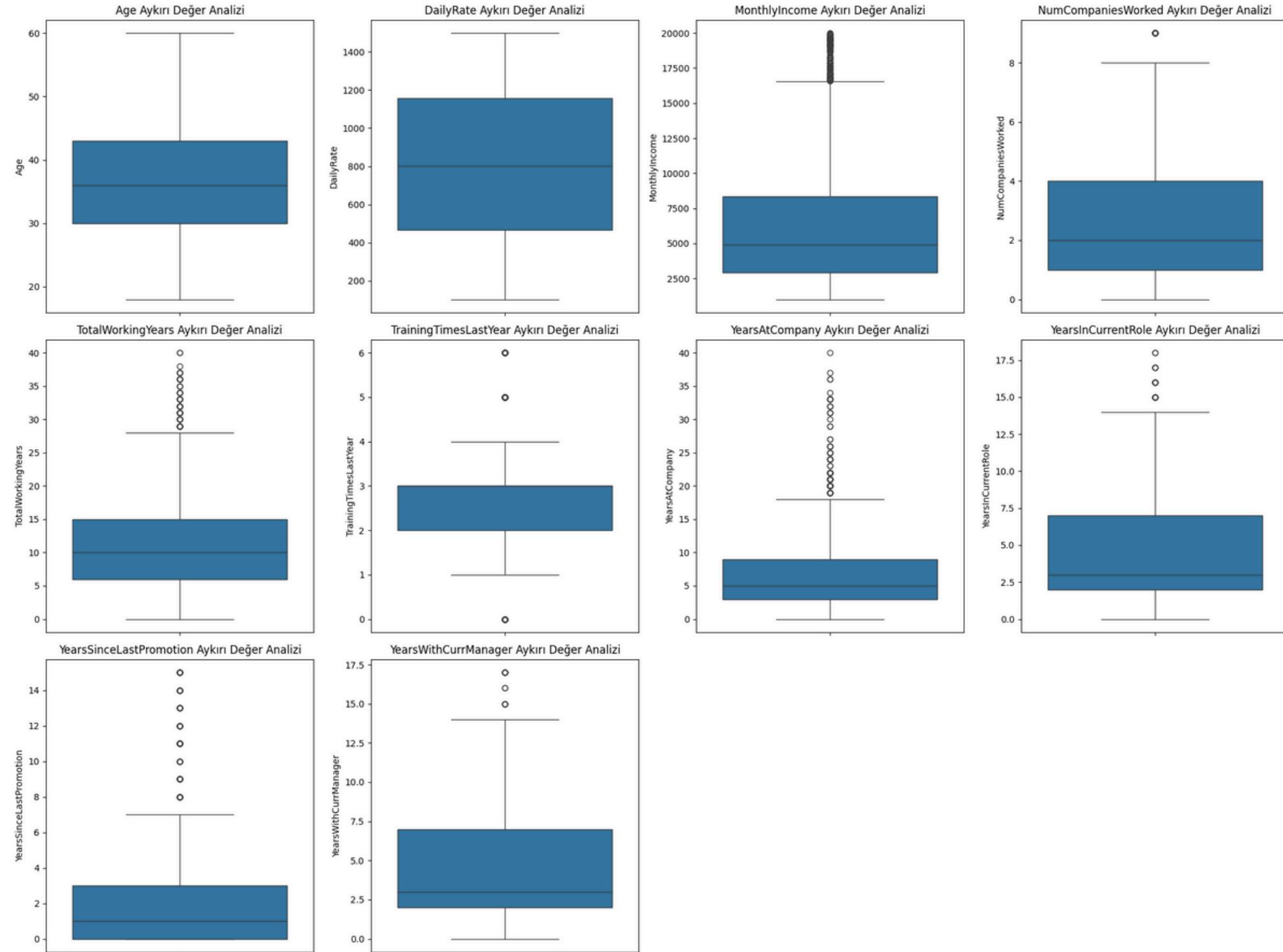
IQR (Interquartile Range), verinin orta %50'sinin yayılımını ölçer.

Aykırı değerler, $Q1 - 1.5 \times IQR$ ve $Q3 + 1.5 \times IQR$ dışındaki değerlerdir.

Aykırı değer analizi IQR yöntemiyle yapıldı; alt ve üst sınırların dışındaki değerler aykırı olarak belirlendi



VERİ ANALİZİ





MODEL YAPIM AŞAMALARI

01

Veri Yükleme

pandas.read_csv()

02

Hedef & Özellik Ayırma;

df['Attrition'].map()
df.drop('EmployeeNumber',
axis=1)

07

Değerlendirme

pandas.read_csv()

06

Modelleme

LogisticRegression
RandomForestClassifier

03

Veri Artırma

- SMOTE()
- feature_swap():200
- feature_mix():200
- feature_noise():200

05

Veri Yükleme

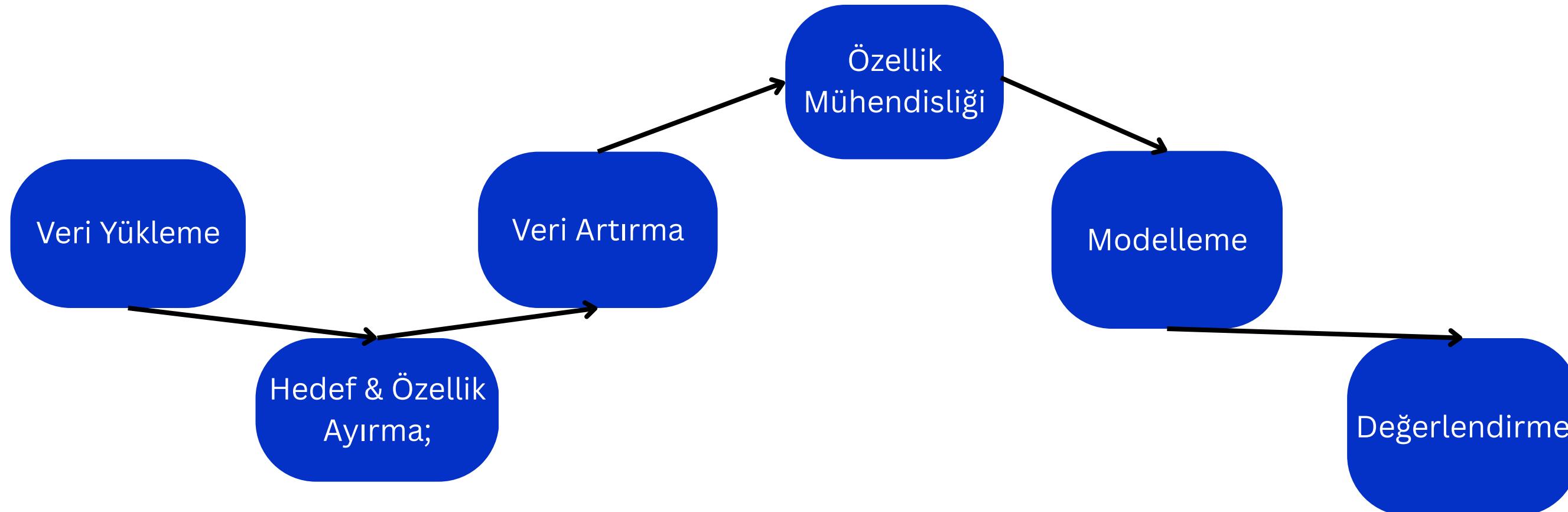
Veri Yükleme;pandas.read_csv()

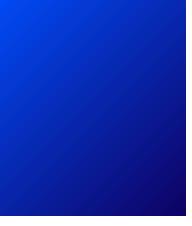
04

Özellik Mühendisliği



YAPIM AŞAMALARI





MODEL

Veri Okuma ve Ön İşleme:

data.csv dosyası yüklandı.

Hedef değişken Attrition sayısallaştırıldı (Yes → 1, No → 0).

Gereksiz sütunlar (ör. EmployeeNumber) çıkarıldı.

Kategorik değişkenler get_dummies() ile one-hot encode edildi.

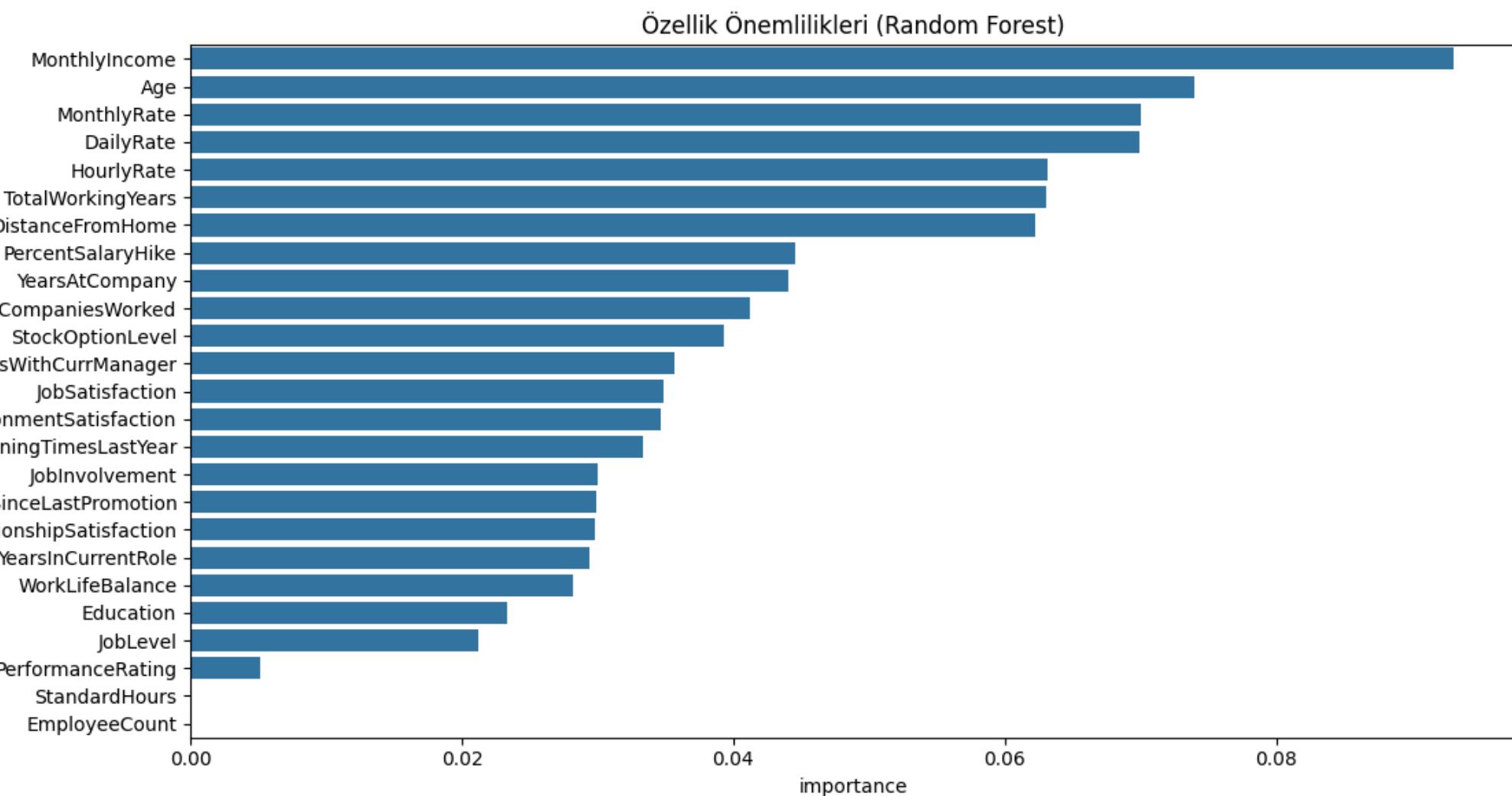
Veri Dengesizliği Sorunu (Imbalanced Data):

SMOTE yöntemiyle azınlık sınıfı (işten ayrılanlar) oversample edildi.

Yapısal Veri Artırma (Structural Data Augmentation):

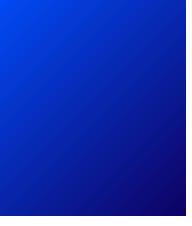
feature_swap, feature_mix, feature_noise teknikleri ile sınıf 1'e ait sentetik örnekler oluşturuldu.

Yeni örneklerle veri seti %60 oranında artırıldı.



Yeni Özellikler (Feature Engineering):

İş tatmini ve çalışma dengesi skorları birleştirilerek yeni değişkenler üretildi (ör. WorkLifeBalanceScore, PromotionSpeed, TenureRatio).



MODEL

Veri Dönüşümü (Preprocessing Pipelines):

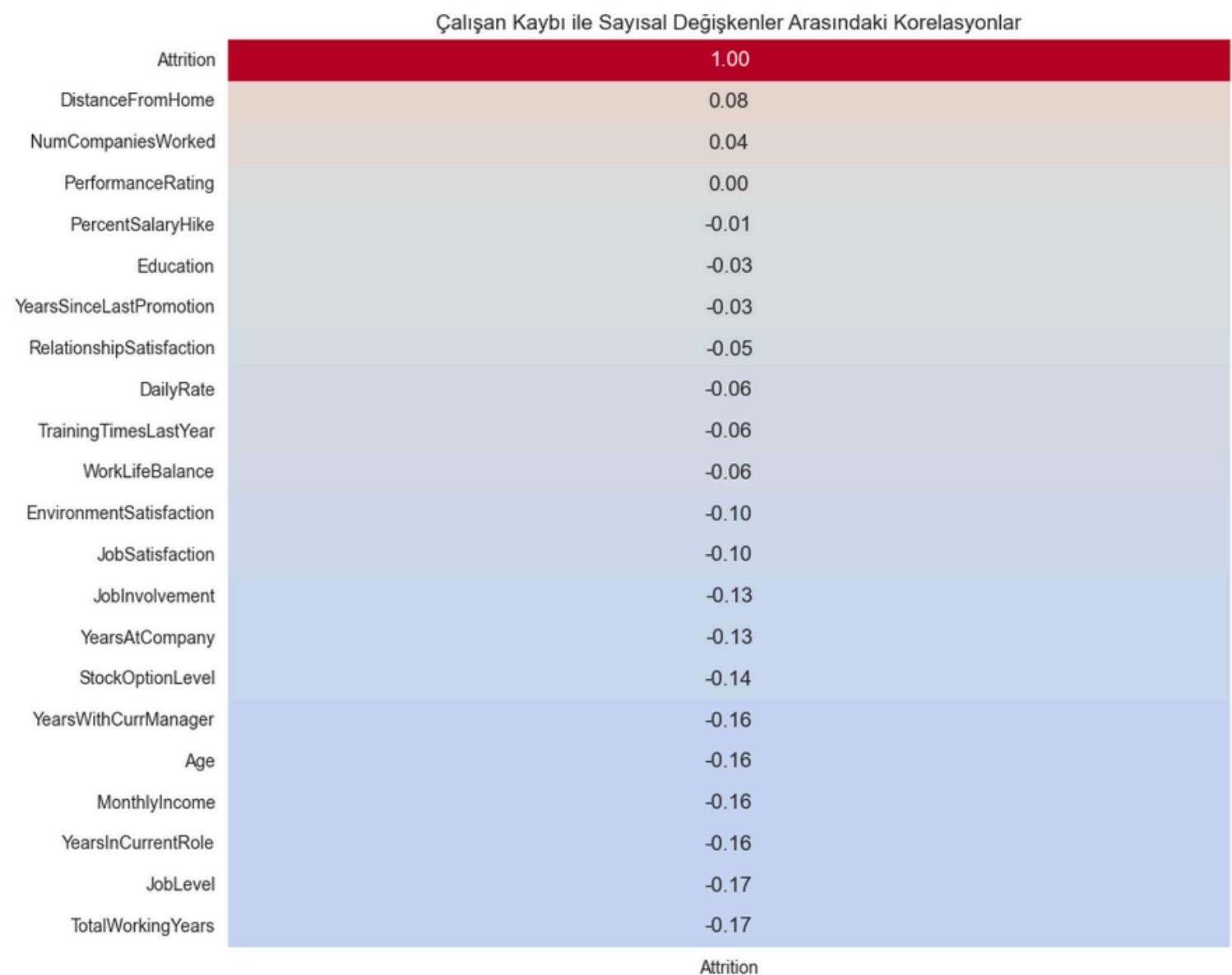
- Sayısal değişkenler: Medyanla doldurma + StandardScaler
- Kategorik değişkenler: En sık değerle doldurma + OneHotEncoder
- ColumnTransformer ile her iki grup tek potada birleştirildi.

Model Kurulumu (Pipeline):

- 3 farklı model pipeline'sı kuruldu:
 - L1 Logistic Regression
 - L2 Logistic Regression
 - Random Forest Classifier

Hiperparametre Optimizasyonu:

- GridSearchCV kullanılarak her model için parametre izgarası tanımlandı.
- En iyi parametre kombinasyonu seçildi ($cv=5$, $scoring='f1'$).

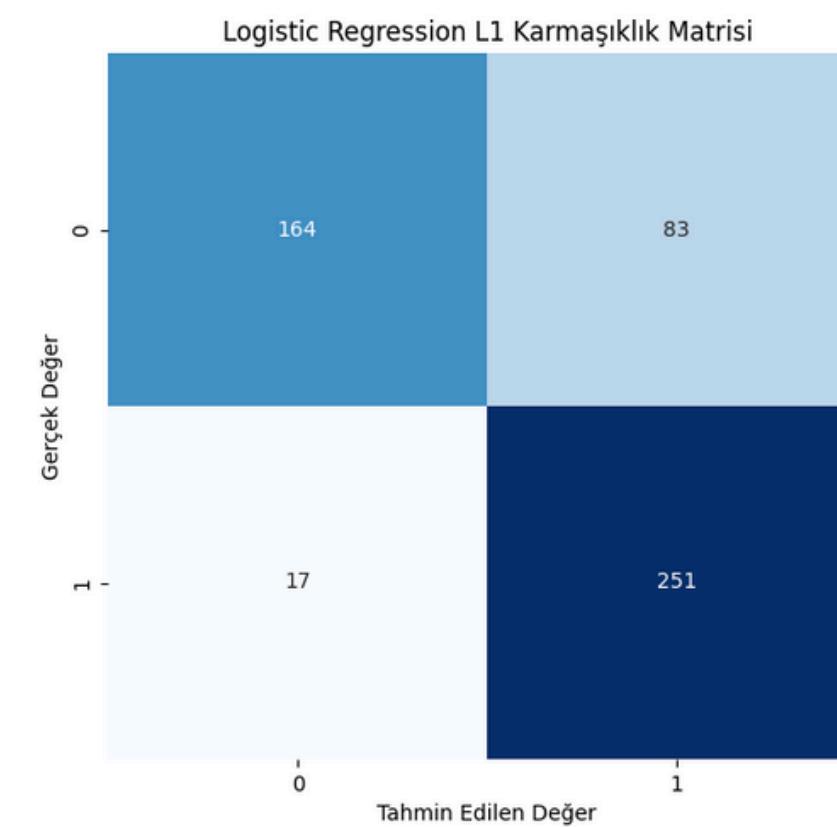


Sonraki →

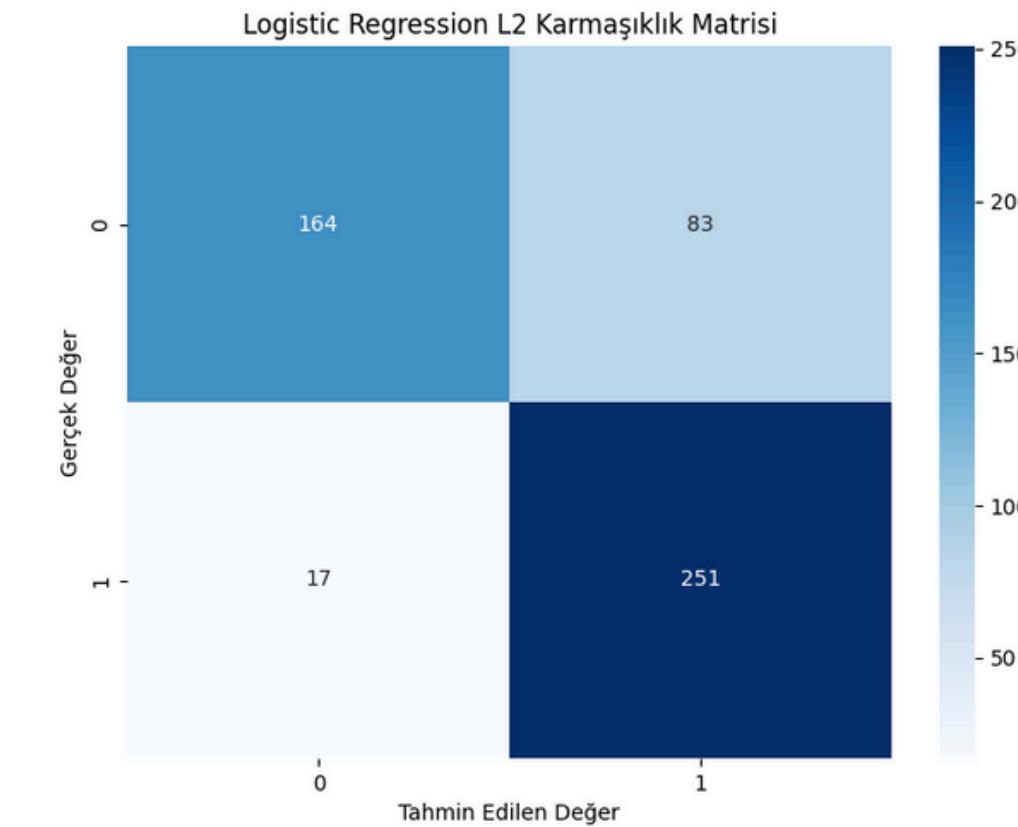


SONUÇLAR

Logical Regression L1 Karmaşıklık Matrisi



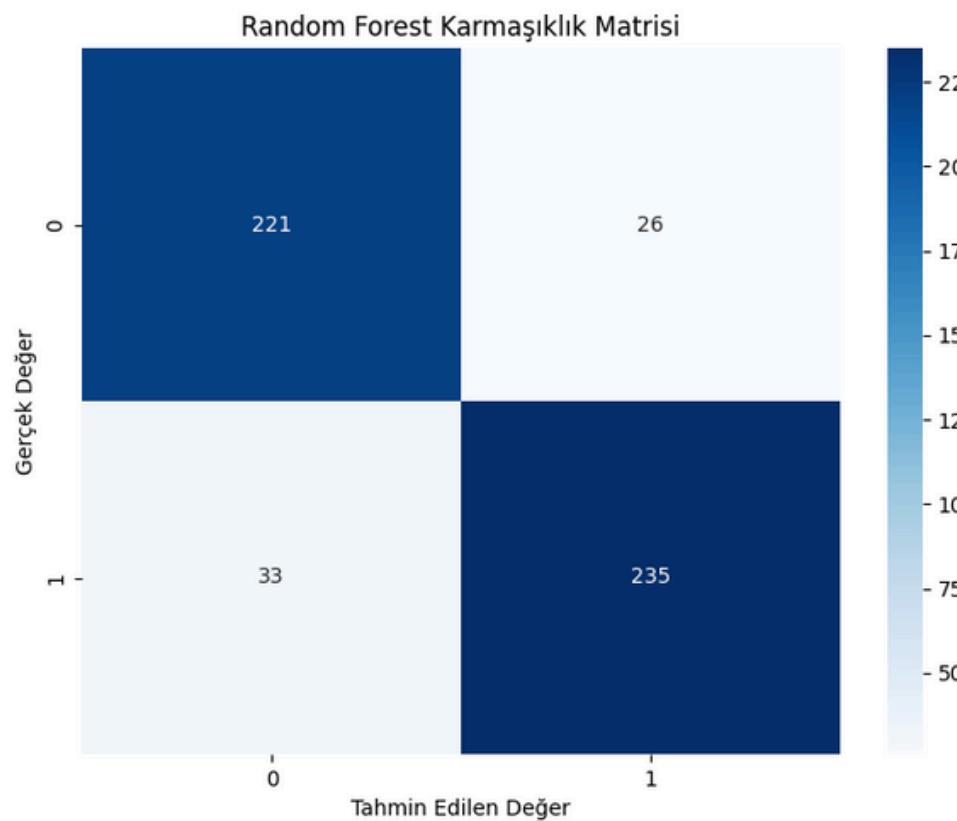
Logical Regression L2 Karmaşıklık Matrisi





SONUÇLAR

Random Forest Karışıklık Matrisi

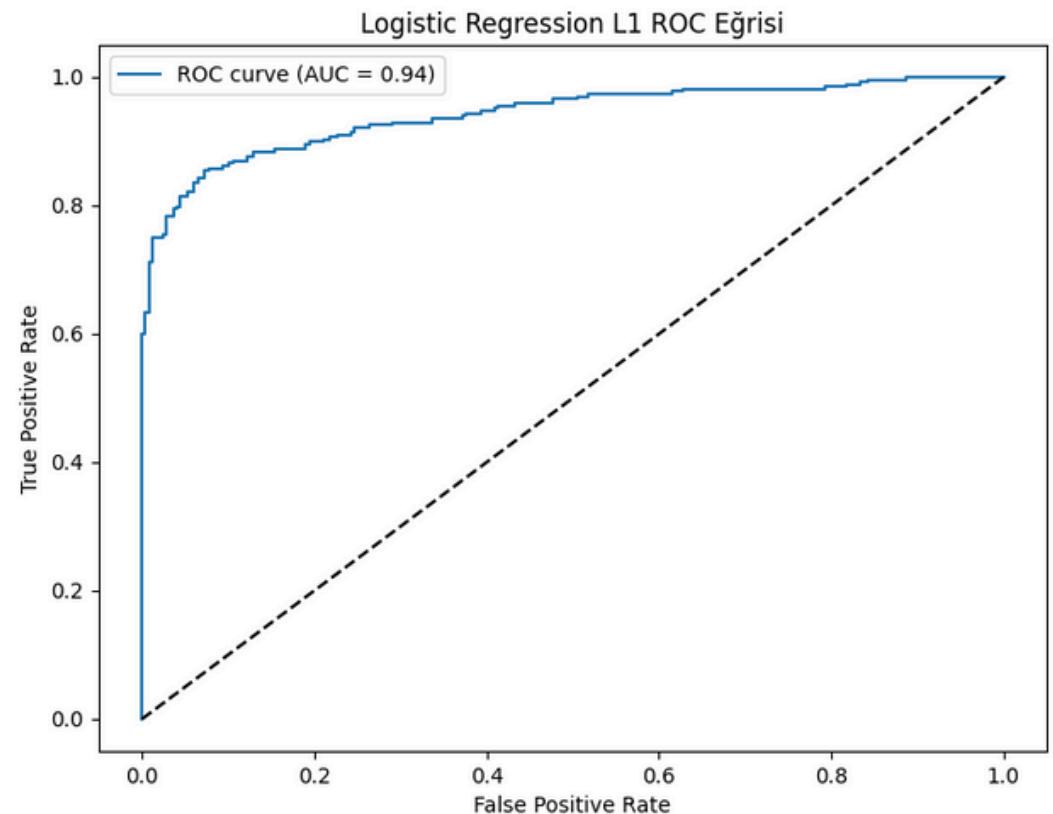


Sonraki →

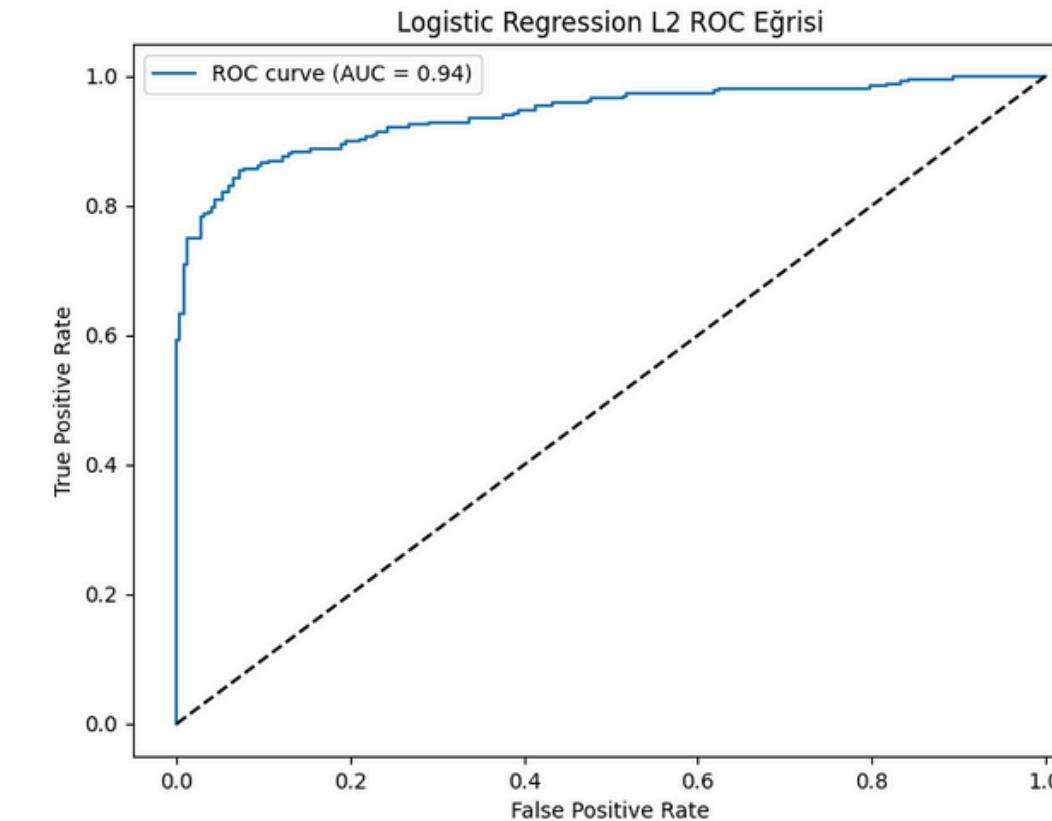


SONUÇLAR

Logical Regression L1 ROC Eğrisi



Logical Regression L2 ROC Eğrisi



Logistic Regression L1 Modeli Hiperparametre Optimizasyonu ve Değerlendirmesi

Fitting 5 folds for each of 18 candidates, totalling 90 fits

Sınıflandırma Raporu:

	precision	recall	f1-score	support
0	0.91	0.66	0.77	247
1	0.75	0.94	0.83	268
accuracy			0.81	515
macro avg	0.83	0.80	0.80	515
weighted avg	0.83	0.81	0.80	515

Logistic Regression L2 Modeli Hiperparametre Optimizasyonu ve Değerlendirmesi

Fitting 5 folds for each of 18 candidates, totalling 90 fits

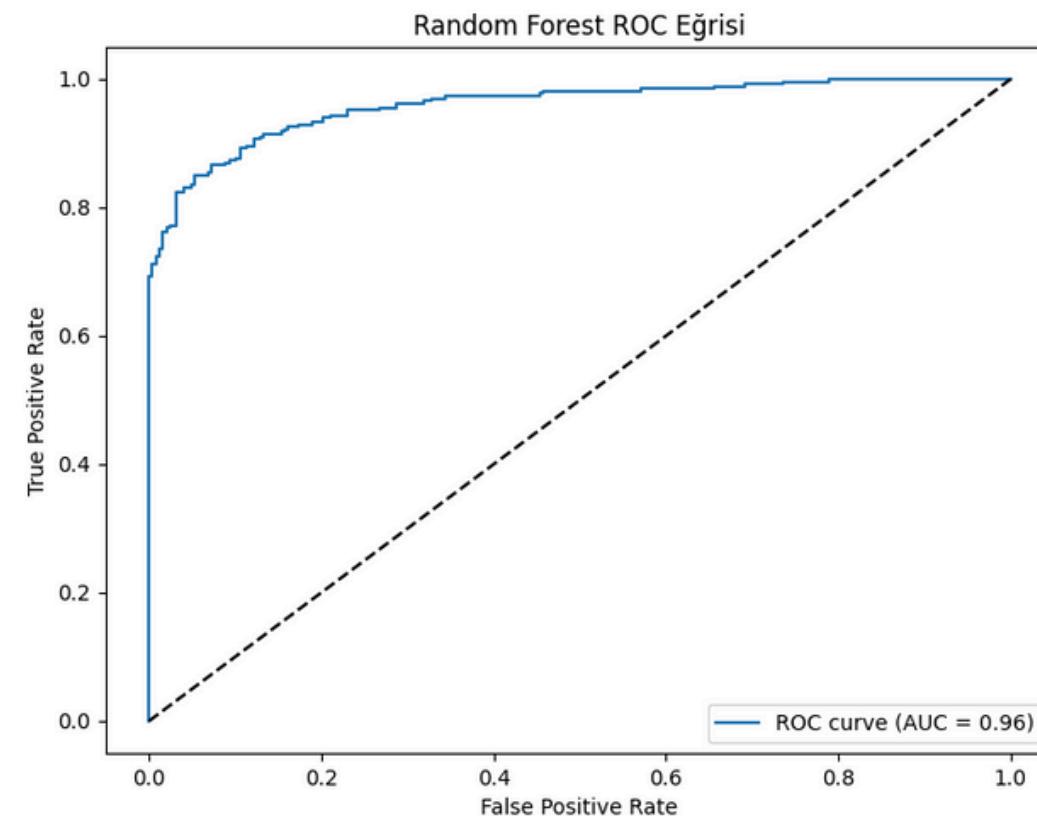
Sınıflandırma Raporu:

	precision	recall	f1-score	support
0	0.91	0.66	0.77	247
1	0.75	0.94	0.83	268
accuracy			0.81	515
macro avg	0.83	0.80	0.80	515
weighted avg	0.83	0.81	0.80	515



SONUÇLAR

Random Forest Roc Eğrisi



Random Forest Modeli Hiperparametre Optimizasyonu ve Değerlendirmesi

```
Fitting 5 folds for each of 216 candidates, totalling 1080 fits
```

Sınıflandırma Raporu:

	precision	recall	f1-score	support
0	0.87	0.89	0.88	247
1	0.90	0.88	0.89	268
accuracy			0.89	515
macro avg	0.89	0.89	0.89	515
weighted avg	0.89	0.89	0.89	515



MODEL KARŞILAŞTIRMASI

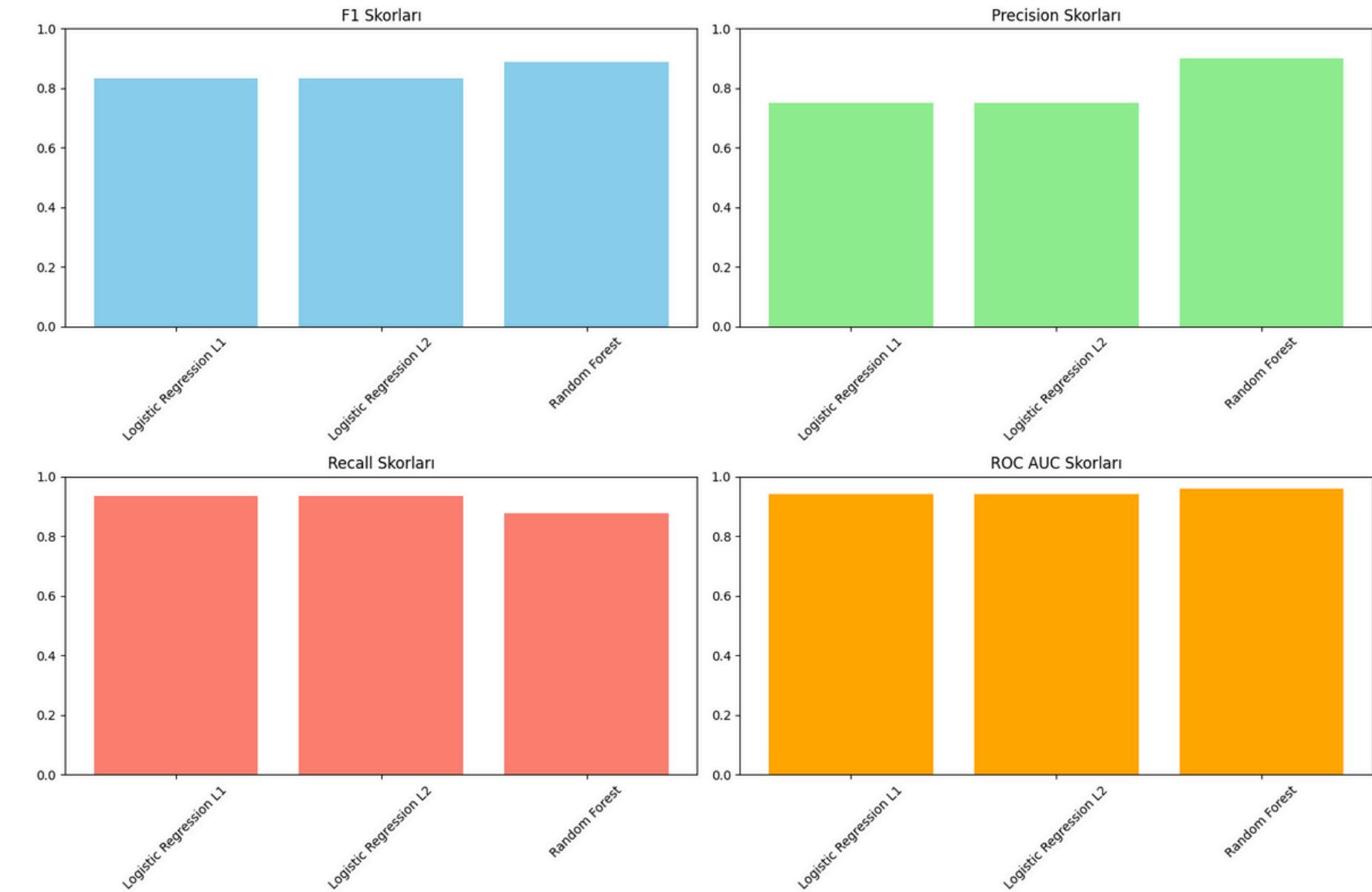
Model Karşılaştırması

Model Performans Karşılaştırması

	Accuracy
Logistic Regression L1	0.805825
Logistic Regression L2	0.805825
Random Forest	0.885437

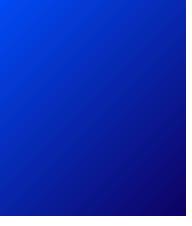
[3 rows x 6 columns]

En İyi Model: Random Forest
En İyi F1 Skoru: 0.888



Hassasiyet Oranı Üzerinden yapılmıştır

Sonraki →



İKİNCİ ADIM; YAPILACAK TESPİTİ



PROBLEM

Araç finansmanı başvurularında müşterilerin ihtiyaç duyduğu bilgilere ulaşamaması, hatalı veri girişi yapması ve süreçte yaşanan zaman kayıpları hem müşteri memnuniyetini düşürmekte hem de operasyonel maliyetleri artırmaktadır.



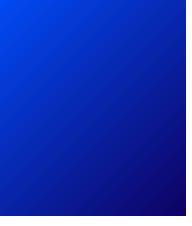
HEDEF / AMAÇ

Müşterilerin yeni veya ikinci el taşit finansmanı ön başvurularını dijital olarak tamamlayabilmeleri için, tüm süreci yöneten ve gerekli bilgileri doğrulayan bir chatbot tasarlamak; ayrıca sık sorulan soruları cevaplayarak kullanıcı deneyimini artırmak.

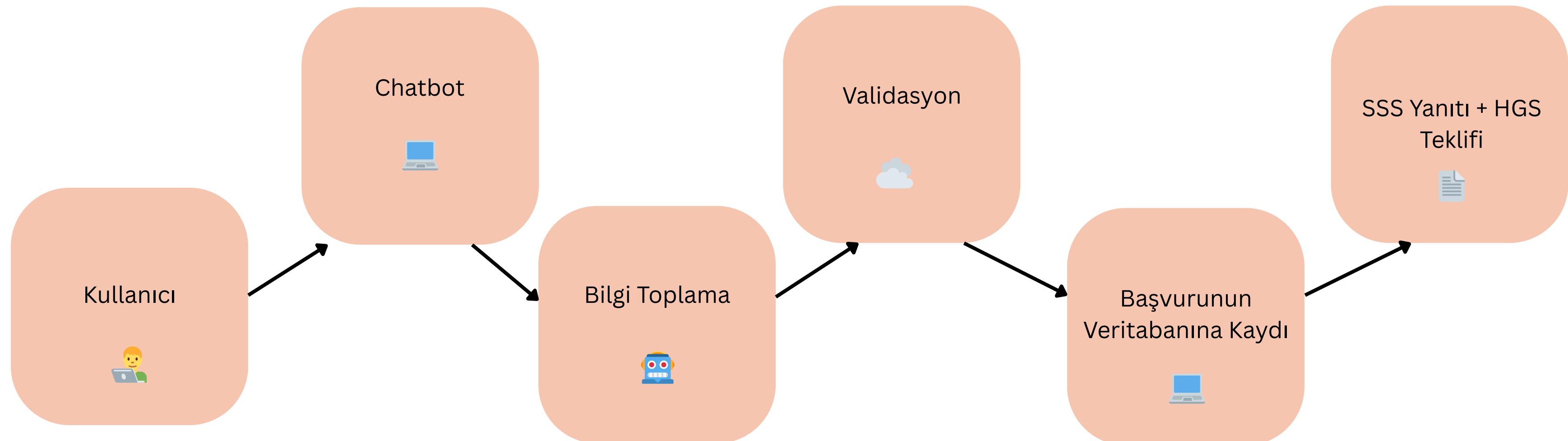


ÇÖZÜM

Chatbot, kullanıcından araç türüne göre gerekli bilgileri adım adım alır, kurallara göre doğrulama yapar, geçerli verilerle başvuru kaydını oluşturur ve HGS teklifi sunar. Sık sorulan sorular chatbot içinde LLM kullanılmadan yanıtlanır, yalnızca gerekiğinde LLM devreye girer.

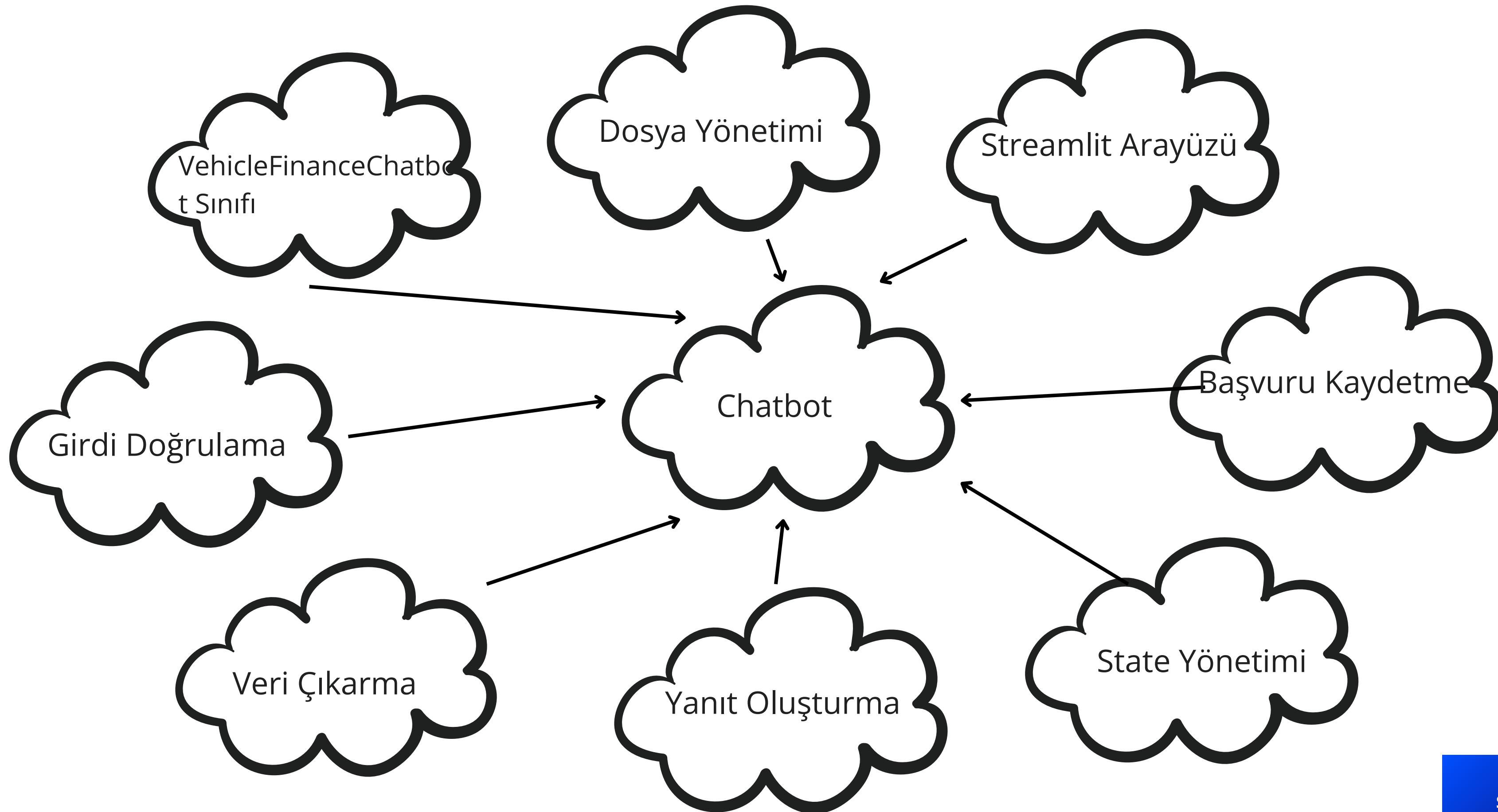


İKİNCİ VAKA İÇİN ÇÖZÜM PLANI





MODEL TASARIMI





İKİNCİ VAKA İÇİN ÇÖZÜM PLANI

VehicleFinanceChatbot Sınıfı

Chatbot'un tüm mantığını ve state yönetimini barındıran ana sınıfıtır.

a. Yapılandırma ve Başlatma

genai.configure(api_key=...): Gemini API'yi başlatır.

chatbot_config.json: Finansman kuralları ve SSS'leri saklar (örneğin, "5 yaş üzeri ikinci el araç alınmaz" gibi).

b. Sistem Promptu: get_system_prompt()

Chatbot'a "bir banka danışmanı gibi" davranışını sağlayan sistem mesajı üretir.

Türkçe, samimi ve rehberlik edici bir tonla çalışır.

Faiz oranı, araç yaşı, kefil gereksinimi gibi tüm kurallar prompt içinde tanımlıdır.

Girdi Doğrulama: validate_data()

Girilen verilerin kural ihlali içerip içermediğini kontrol eder.

Veri Çıkarma: extract_info_from_text()

Kullanıcı mesajından sayı, TCKN veya metin gibi bilgiler çıkarır. Regex temelliidir.

State Yönetimi: process_message()

Tüm mesaj akışını ve adım kontrolünü yöneten ana fonksiyondur.

Adımlar:

greeting → determine_type → data_collection → confirmation → hgs_offer → end



CHATBOT GELİŞTİRME ADIMLARI

İhtiyaç Analizi ve Kapsam Belirleme

Use-case'in netleştirilmesi: Araç finansmanı başvurusu için bir LLM destekli yardımcı. Gerekli alanlar ve iş kuralları belirlendi (örneğin yaşı sınırı, kredi oranı).

Veri Yapısının Planlanması

JSON tabanlı yapı kurularak iş kuralları (finance_rules), mesajlar, SSS, validasyonlar ayrıldı.

LLM Prompt Mimarisi Tasarımı

Rol tanımı, dil ve yanıt tarzı gibi davranış kuralları system_prompt ile tanımlandı.

Sohbet Akışının (Workflow) Kodlanması

Her adım için self.current_step ile state yönetimi yapıldı (örnek: greeting, confirmation).

Doğrulama ve Veri Girişi Yönetimi

TCKN, tutar gibi bilgiler validate_data() ve extract_info_from_text() fonksiyonlarıyla denetlendi.

Uygulama ve UI Geliştirme

Streamlit ile kullanıcı dostu frontend geliştirildi. CSS ile zenginleştirilmiş görsel deneyim sağlandı.

Veri Saklama ve Başvuru Kaydı

Başvurular JSON dosyasına kaydedildi (save_application()).



İKİNCİ VAKA İÇİN ÇÖZÜM PLANI

Streamlit Arayüzü

1. st.set_page_config ve CSS

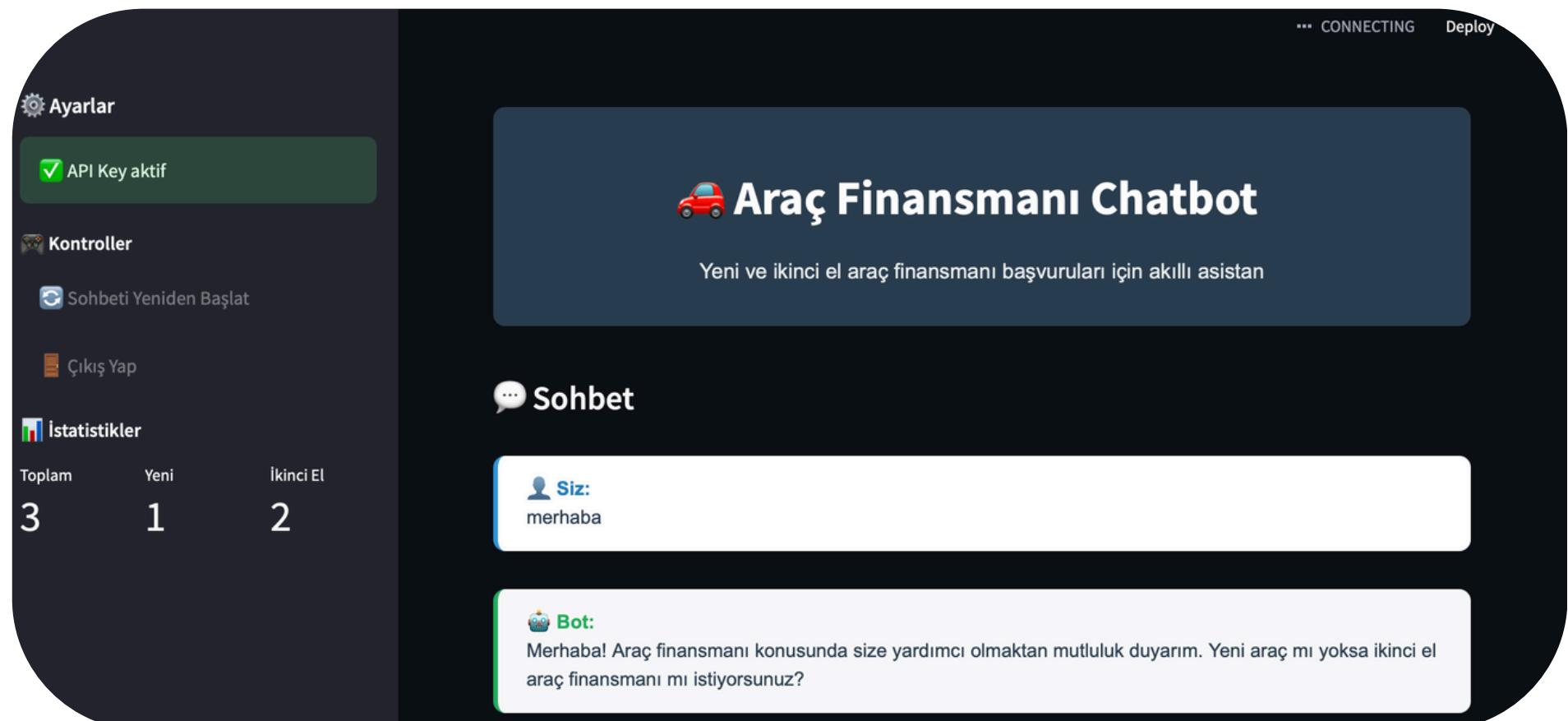
- Geniş ekran, responsive, sade ve kurumsal görünümlü arayüz.
- Kullanıcı ve bot mesajları renkli kutularda gösterilir.

2. main() Fonksiyonu

- API Key doğrulama → Chat başlatma
- Sohbet geçmişini yönetimi
- Yanıtları st.chat_input() ve özel HTML ile gösterir.

3. st.sidebar

- API Key girişi ve validasyonu
- Sohbet sıfırlama ve çıkış
- Toplam başvuru sayısı gibi özet metrikler



Sonraki →



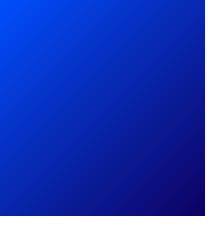
CHAT BOT ARAYÜZÜ



Araç Finansmanı Chatbot

Yeni ve ikinci el araç finansmanı başvuruları için akıllı asistan

Sonraki →



**DİNLEDİĞİNİZ
İÇİN
TEŞEKKÜRLER**

