



LightGBM & Streamlit

Dilan Nihadioğlu



Ben Kimim ?

Merhabalar, ben Dilan Nihadioğlu. Aydın Adnan Menderes Üniversitesi'nde Bilgisayar Mühendisliği bölümündeyim. LightGBM ile ilgili yaptığım projeleri streamlit ile sunacağım.



BACK

NEXT

Sunum Planı



LightGBM



Müşteri Satın Alma
Analizi



Gelir Tahmini



Göğüs Kanseri



Soru & Cevap

BACK



NEXT

LightGBM Nedir?

LightGBM, **büyük veri** setleriyle **hızlı ve etkili** bir şekilde çalışabilen, **karar ağaçları** üzerine kurulu bir makine öğrenimi algoritmasıdır. Sınıflandırma ve regresyon gibi görevlerde güçlü performans gösterir. LightGBM, özellikle hız ve doğruluk gerektiren projelerde öne çıkar.

BACK



NEXT

LightGBM Neden Popüler?

LightGBM, **hesaplama gücü ve daha hızlı sonuç vermesi** nedeniyle "Light" olarak adlandırılır. Çalıştırmak için **daha az bellek** gerektirir ve **büyük miktarda veriyle** başa çıkabilir.

BACK



NEXT



LightGBM Algoritması Ne Yapar?

Eksik Değerlerle Otomatik İşleme: LightGBM, veri setinizde eksik değerler varsa otomatik olarak bu değerleri işleyebilir.

Histogram Tabanlı Bölme: Karar ağaçlarında, verileri bölmek için kullanılan noktalar (yani hangi değere göre verilerin ikiye ayrılacağı) LightGBM'de histogram yöntemiyle bulunur. Bu yöntem, veriyi daha hızlı ve verimli şekilde böler.

Gradyan Tabanlı Örnekleme (GOSS): LightGBM, verinin sadece önemli kısımlarını kullanarak (gradyanlara göre) daha az veriyle daha iyi sonuçlar elde edebilir. Bu da eğitim süresini kısaltır.

Yaprak Bazında Ağaç Büyümesi: Daha doğru tahminler için ağaç en etkili yapraklardan büyütür.

Özel Özellik Paketi Tekniği: Veriyi sıkıştırarak daha az bellek kullanır, hızlı çalışır.

[BACK](#)[NEXT](#)

Müşteri Satın Alma Analizi

Müşteri Satın Alma Analizi

Bu projede, müşterinin satın alma kararlarını inceleyeceğimiz bir sınıflandırma problemini ele alacağız.

Show dataframe

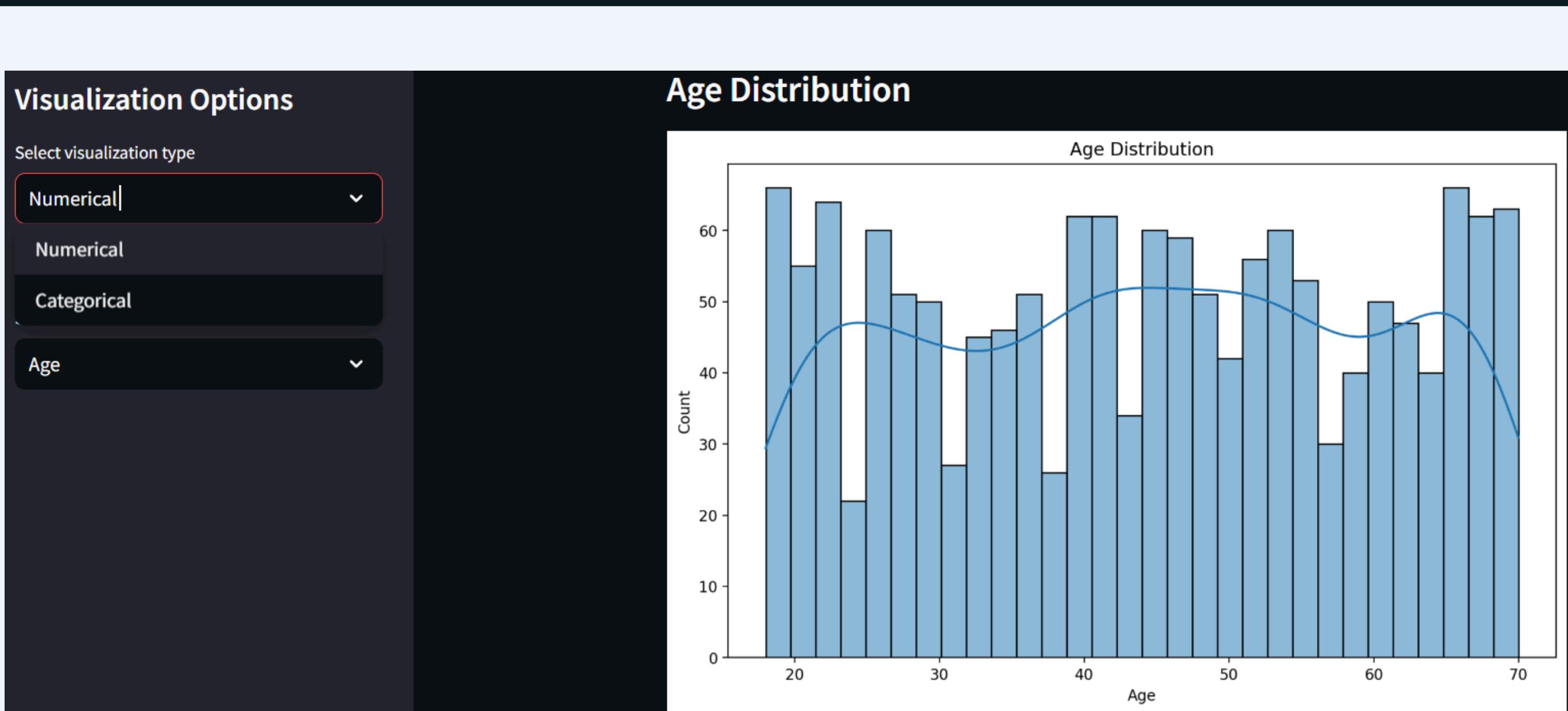
	Age	Gender	AnnualIncome	NumberOfPurchases	ProductCategory	TimeSpentOnWebsite	LoyaltyScore
0	40	1	66,120.2679	8	0	30.5686	
1	20	1	23,579.7736	4	2	38.2401	
2	27	1	127,821.3064	11	2	31.6332	
3	24	1	137,798.6231	19	3	46.1671	
4	31	1	99,300.9642	19	1	19.8236	
5	66	1	37,758.1175	14	4	17.8275	
6	39	1	126,883.3853	16	3	42.0854	
7	64	1	39,707.3597	13	2	17.1903	
8	43	0	102,797.3013	20	1	6.0235	
9	20	1	63,854.9211	16	0	38.5725	

BACK

NEXT

Show describe statistic

Streamlit ile Proje Detayları



BACK

NEXT



Streamlit ile Proje Detayları

Visualization Options

Select visualization type

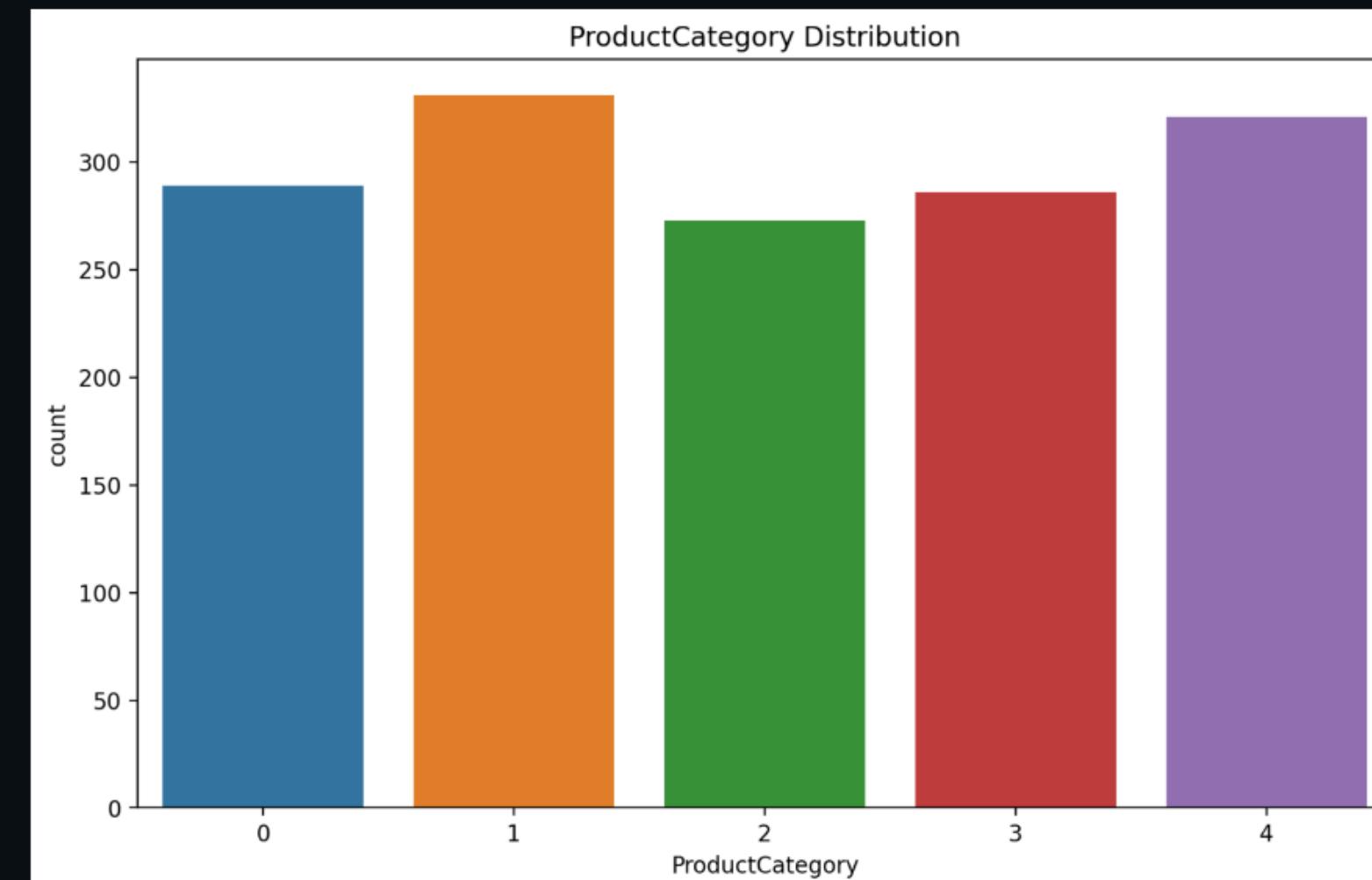
Categorical

Categorical Features

Select categorical feature

ProductCategory

ProductCategory Distribution



BACK

NEXT



Streamlit ile Proje Detayları

Model Comparison Metrics

	Accuracy	Precision	Recall	F1 Score	Execution Time
Random Forest	0.9533	0.9914	0.8984	0.9426	0.2496
Support Vector Machine	0.8700	0.8870	0.7969	0.8395	0.0625
LightGBM	0.9467	0.9828	0.8906	0.9344	0.0869
XGBoost	0.9500	0.9829	0.8984	0.9388	0.0781

Select a model to view its confusion matrix

LightGBM

Random Forest

Support Vector Machine

LightGBM

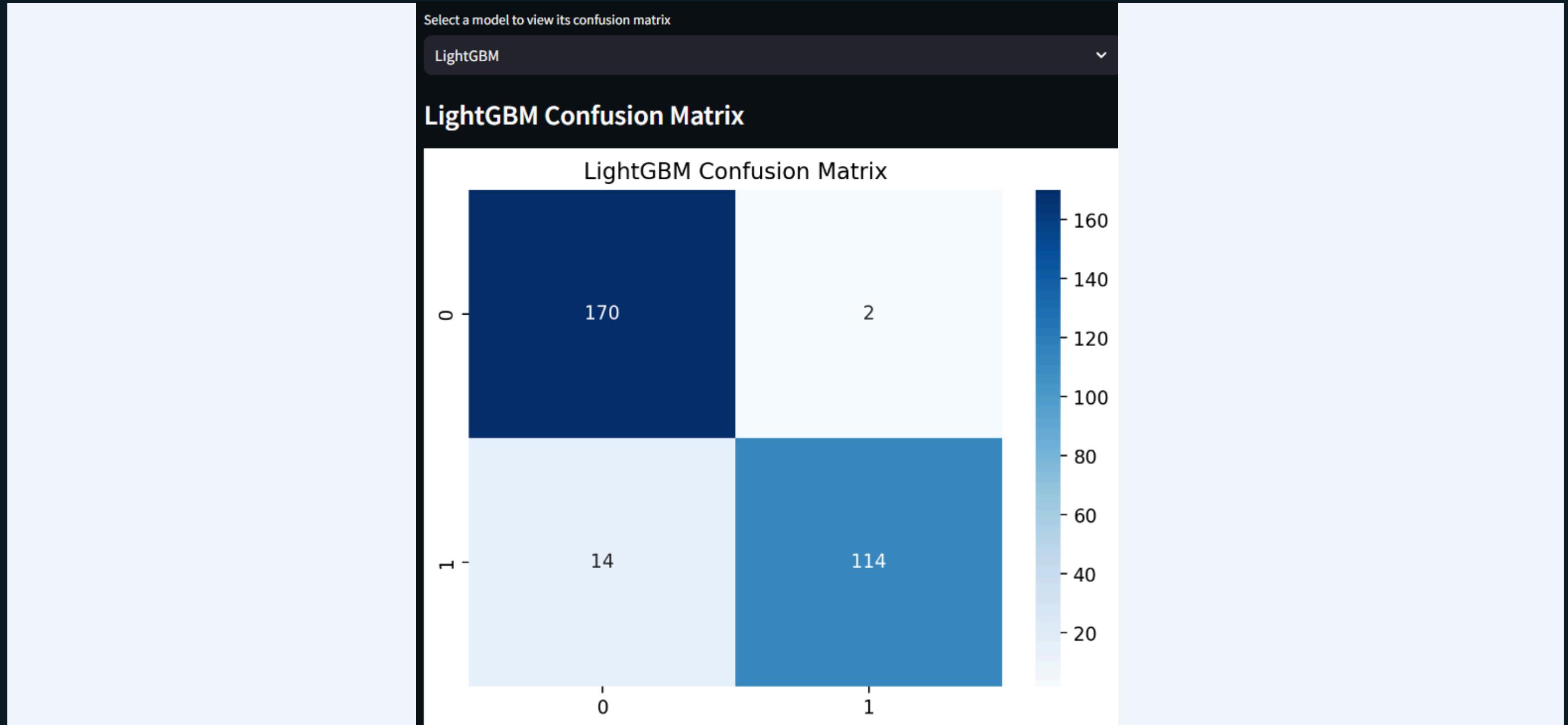
XGBoost

BACK

NEXT



Streamlit ile Proje Detayları





Göğüs Kanseri Tahmini İçin LightGBM Kullanımı

Breast Cancer Prediction with LightGBM ↗

Bu projede, LightGBM modeli ile göğüs kanseri sınıflandırması yapacağız.

Show dataframe

	mean radius	mean texture	mean perimeter	mean area	mean smoothness	mean compactness	me:
0	17.99	10.38	122.8	1,001	0.1184	0.2776	
1	20.57	17.77	132.9	1,326	0.0847	0.0786	
2	19.69	21.25	130	1,203	0.1096	0.1599	
3	11.42	20.38	77.58	386.1	0.1425	0.2839	
4	20.29	14.34	135.1	1,297	0.1003	0.1328	
5	12.45	15.7	82.57	477.1	0.1278	0.17	
6	18.25	19.98	119.6	1,040	0.0946	0.109	
7	13.71	20.83	90.2	577.9	0.1189	0.1645	
8	13	21.82	87.5	519.8	0.1273	0.1932	
9	12.46	24.04	83.97	475.9	0.1186	0.2396	

BACK

NEXT

Görselleştirme Seçenekleri

Sayısal Kolon Seçin

mean radius

Grafik Tipi Seçin

Violin Plot

Histogram

Violin Plot

Box Plot

Streamlit ile Proje Detayları

Görselleştirme Seçenekleri

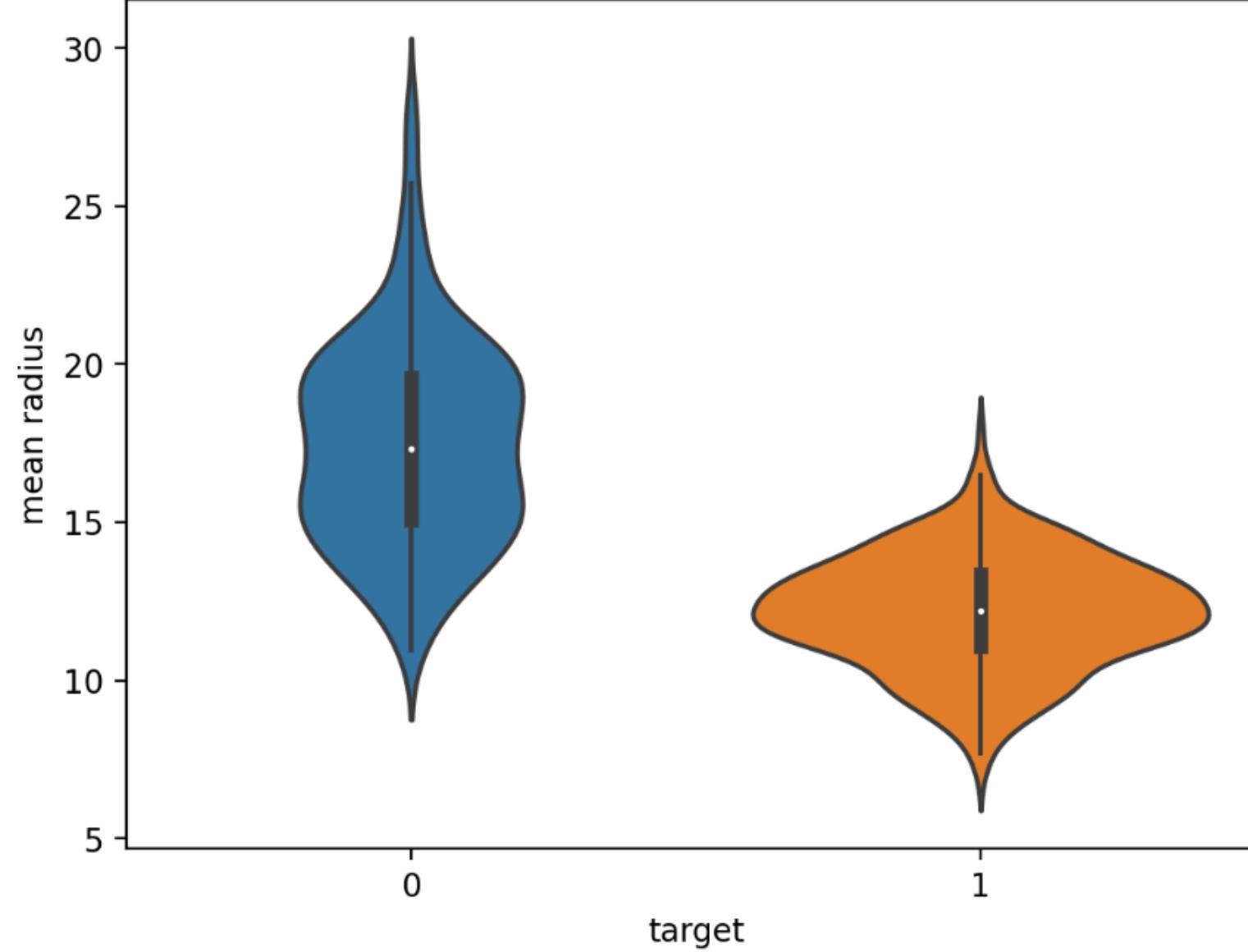
Sayısal Kolon Seçin

mean radius

Grafik Tipi Seçin

Violin Plot

Violin Plot of mean radius



BACK

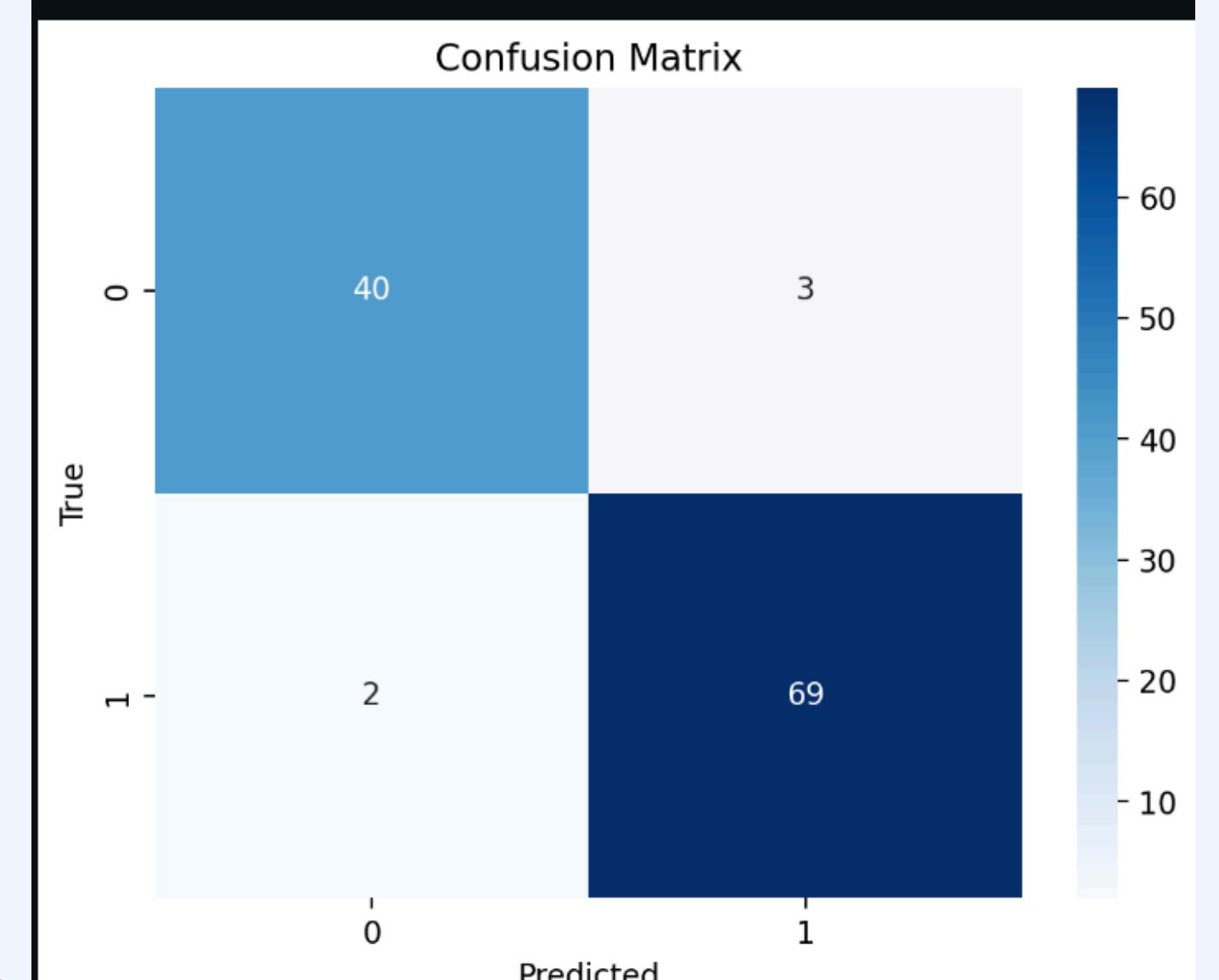
NEXT

Streamlit ile Proje Detayları

Model Performance Comparison

	Model	Accuracy	Precision	Recall	F1-Score	Time (s)
0	LightGBM	0.9561	0.9583	0.9718	0.965	0.125
1	Random Forest	0.9649	0.9589	0.9859	0.9722	0.25
2	XGBoost	0.9561	0.9583	0.9718	0.965	0.2028

Confusion Matrix



BACK

NEXT

LightGBM ile Gelir Tahmini

Adult-Income

Bu proje, Adult Income veri setini kullanarak bireylerin yıllık gelirinin 50K'nın üzerinde mi yoksa altında mı olduğunu tahmin etmeyi amaçlar.

Show dataframe

Show describe statistic

	count	mean	std	min	25%	50%	75%	max
age	45,222	38.5479	13.2179	17	28	37	47	90
fnlwgt	45,222	189,734.7343	105,639.1951	13,492	,388.25	178,316	237,926	1,490,400
educational-num	45,222	10.1185	2.5529	1	9	10	13	16
capital-gain	45,222	1,101.4303	7,506.4301	0	0	0	0	99,999
capital-loss	45,222	88.5954	404.9561	0	0	0	0	4,356
hours-per-week	45,222	40.938	12.0075	1	40	40	45	99

BACK



NEXT

Streamlit ile Proje Detayları

Dataset Visualization

Bar Chart for relationship

relationship	count
Own-child	~6500
Husband	~18000
Not-in-family	~11500
Unmarried	~4500
Wife	~2000
Other-relative	~1500

Categorical Data Visualization

Select categorical column

relationship

Select chart type for categorical data

Bar Chart

Numerical Data Visualization

Select numerical column

age

Select chart type for numerical data

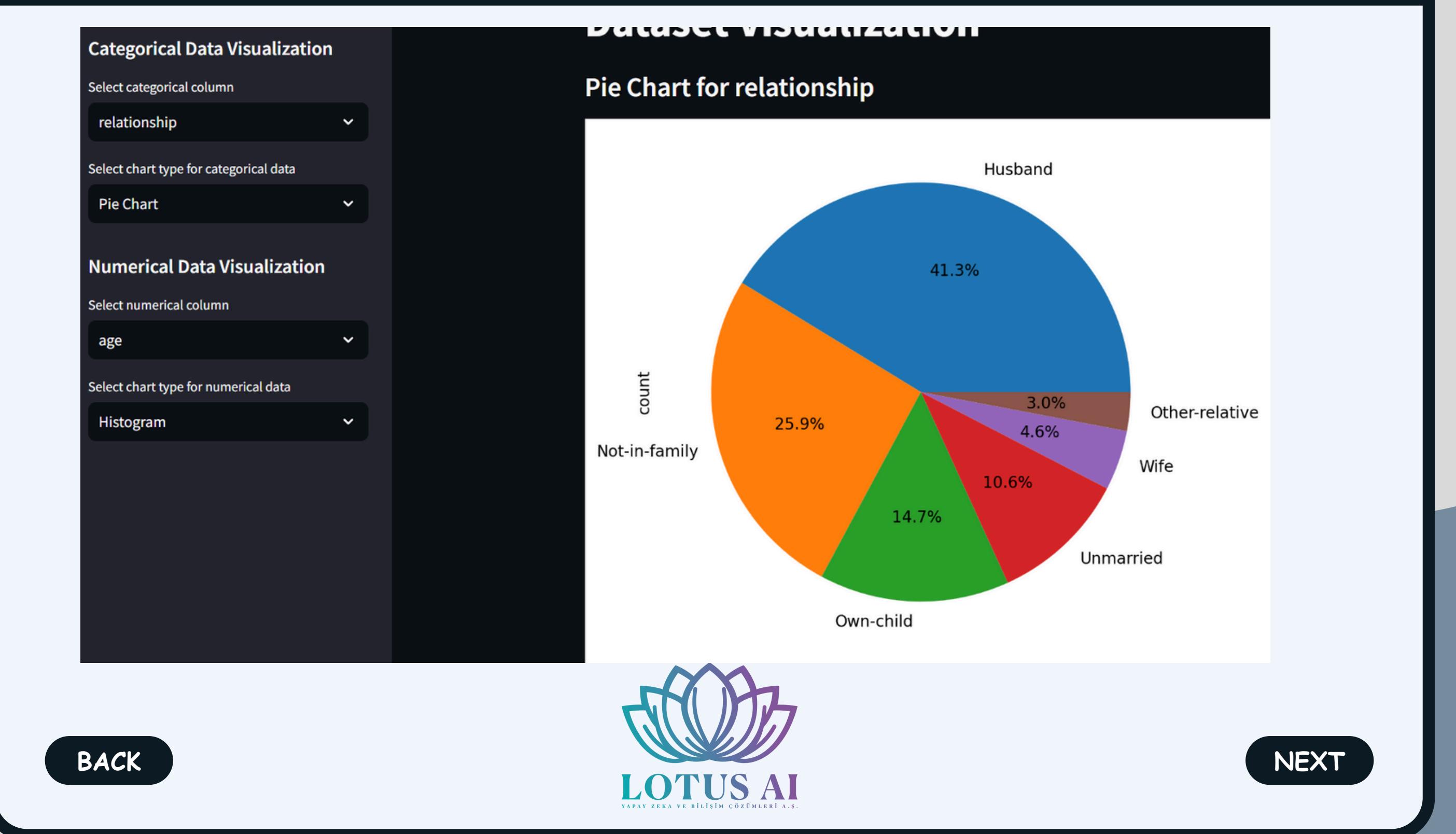
Histogram

BACK

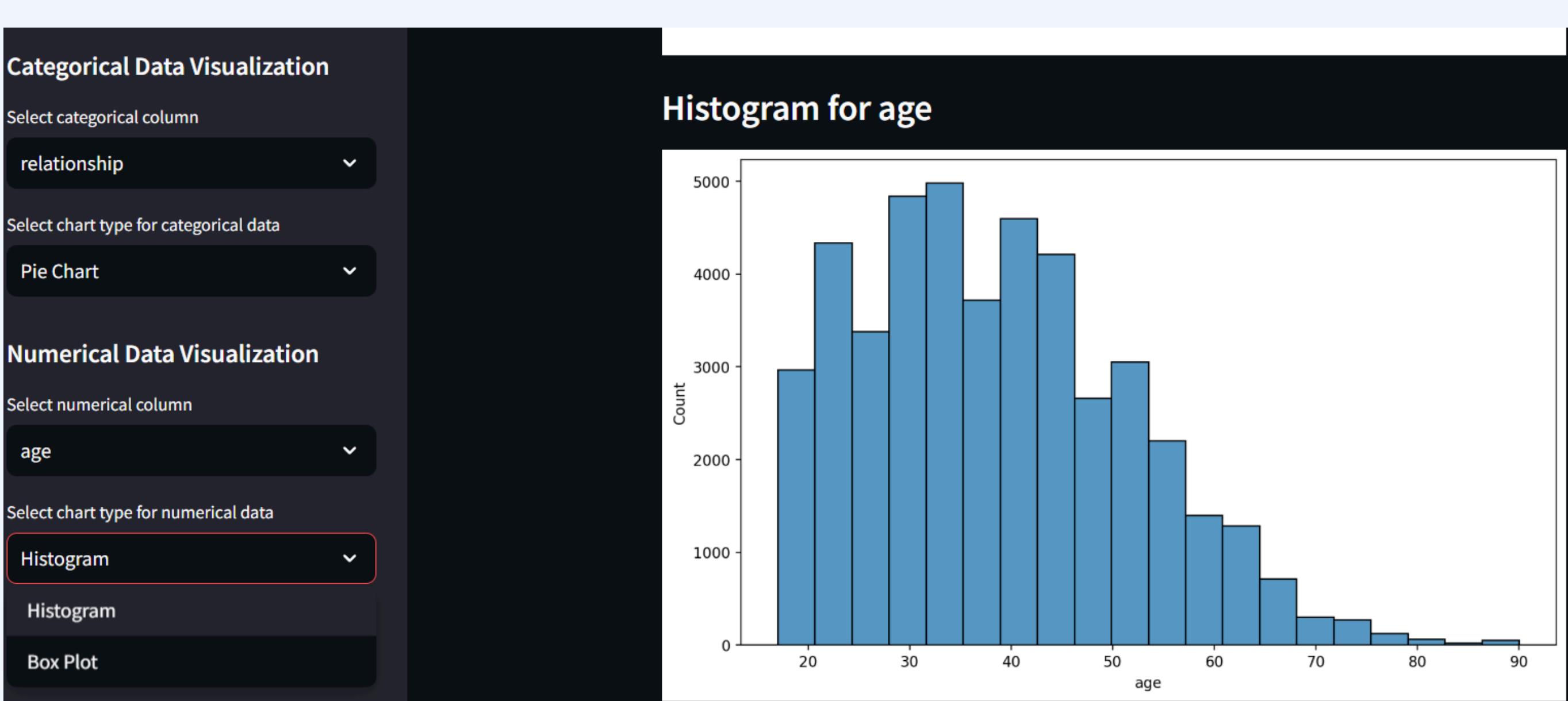
NEXT

LOTUS AI
YAPAY ZEKA VE BİLİŞİM ÇÖZÜMLERİ A.Ş.

Streamlit ile Proje Detayları



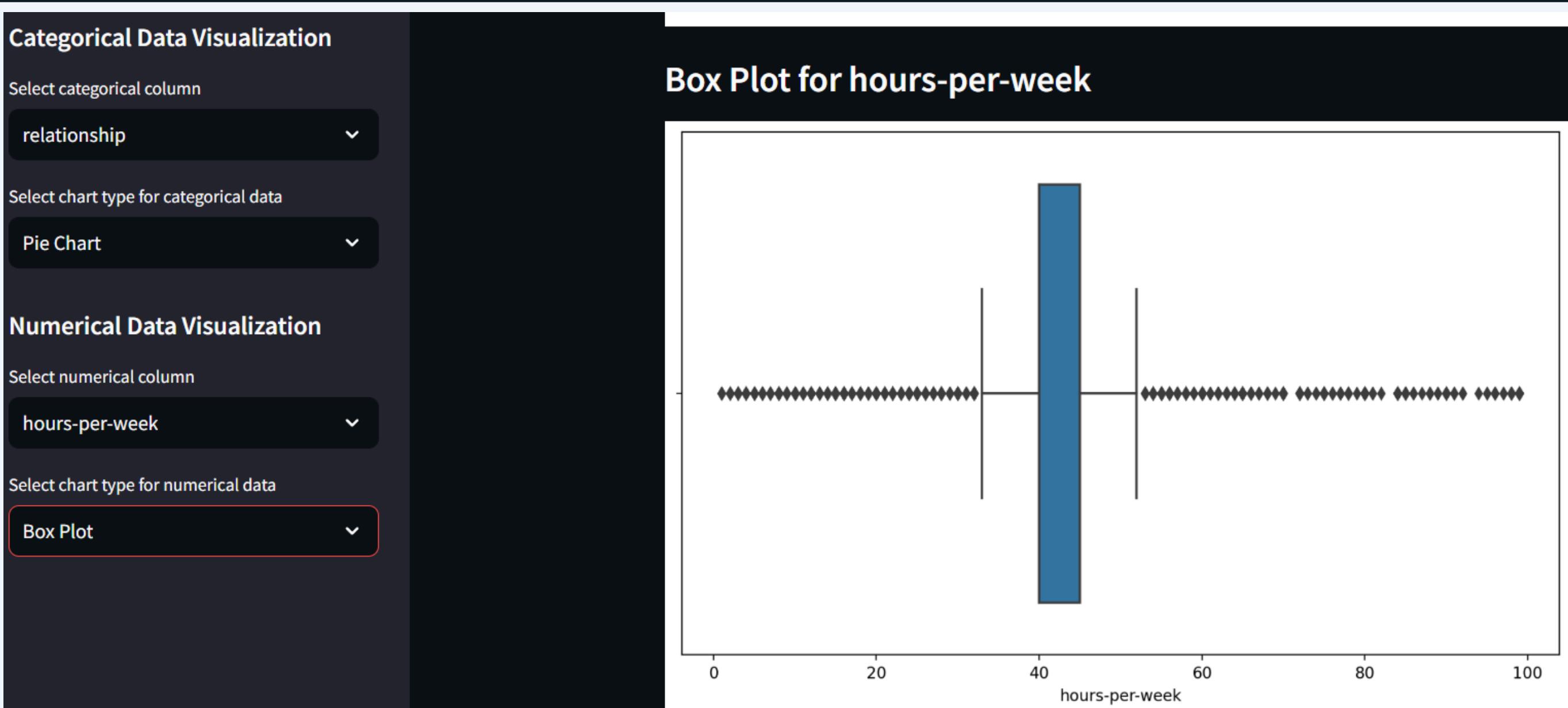
Streamlit ile Proje Detayları



BACK

NEXT

Streamlit ile Proje Detayları



BACK

NEXT

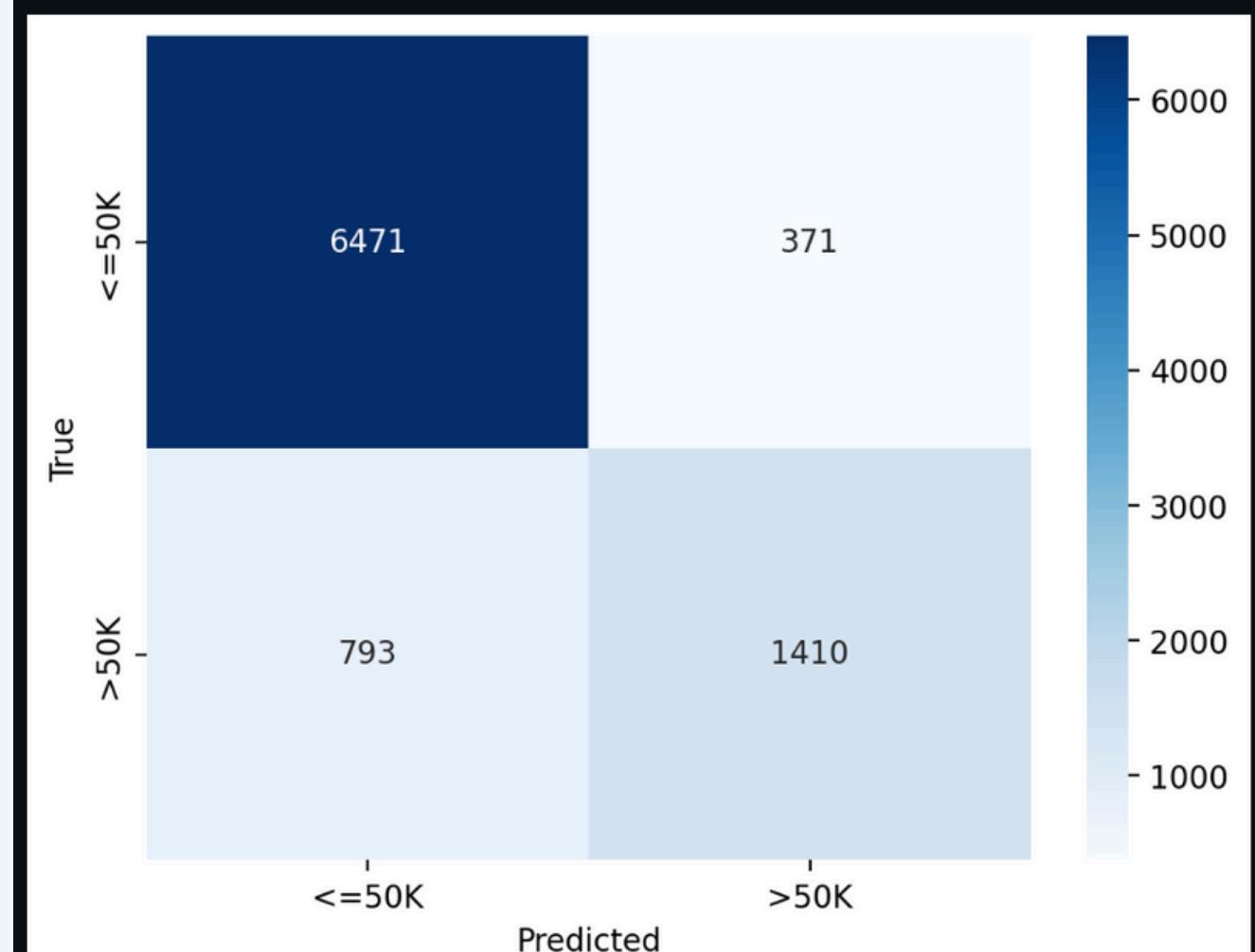


Streamlit ile Proje Detayları

Model Metrics

	Metric	LightGBM	XGBoost	Random Forest
0	Accuracy	0.8713	0.8721	0.8561
1	Log Loss	0.2863	0.2919	0.3811
2	ROC AUC	0.9250	0.9225	0.9047
3	Time (seconds)	0.2347	0.2656	4.9714

Confusion Matrix



BACK



Soru & Cevap



BACK

NEXT

