
ONLINE SHOPPERS PURCHASING INTENTION DATASET

MACHINE LEARNING PROJECT

HAZIRLAYANLAR

- Kadir Erimhan
- Haydar Şensoy
- Ali F. Baynazoğlu

UCI
Machine Learning Repository
Center for Machine Learning and Intelligent Systems

Home Datasets Data Sets Data Set Examples
[Home](#) [Datasets](#) [Data Sets](#) [Data Set Examples](#)
[Home](#) [Datasets](#) [Data Sets](#) [Data Set Examples](#)

Check out the beta version of the new UCI Machine Learning Repository on an internally hosted! Contact us if you have any ideas, questions, or concerns. Click here to get the beta version.

Online Shoppers Purchasing Intention Dataset Data Set

Download [Data Files](#) [Data Set Description](#)

Dataset Size: 10,135 sessions in the dataset. 54,475 (54,422) were negative class samples that did not end with shopping, and the rest (545) were positive class samples ending with shopping.

Data Set Characteristics	Multi-class	Number of Instances	10135	Area	Real-world
Attribute Characteristics	Integer, Real	Number of Attributes	18	Date Created	2016-03-31
Associated Task(s)	Classification, Clustering	Missing Values?	No	Number of Web Sites	23736

Source:

U. Demirel
Department of Computer Engineering, Faculty of Engineering and Natural Sciences, Sakarya University,
54189 Sogutlu, Sakarya, Turkey

A. Yildiz
Verim Information Technologies Consulting and Trade,
44120 Sakarya, Turkey

Data Set Information:

The dataset consists of session records belonging to 10,135 sessions. The dataset was formed so that each session corresponds to a different user. It is a time series data set, meaning it is a specific campaign, specific day, and specific product.

Attribute Information:

The screenshot shows the UCI Machine Learning Repository website. The main heading is "Online Shoppers Purchasing Intention Dataset Data Set". Below this, there are links for "Download Data Files" and "Data Set Description". A table provides details about the dataset, including the number of instances, attributes, and associated tasks. The source is listed as "Source: UCI ML Repository". The data set information section explains that the dataset consists of session vectors belonging to 10,000 sessions, with each session represented by a 10-dimensional vector. The attribute information section is also visible.

Data Set Characteristics	Multi-class	Number of Instances	1000	Attribute	Session
Attribute Characteristics	Integer, Real	Number of Attributes	10	Date Created	2010-03-31
Associated Task(s)	Classification, Clustering	Missing Values?	No	Number of Web Sites	2010-03

Source:
UCI ML Repository
Department of Computer Engineering, Faculty of Engineering and Natural Sciences, Sakarya University, 54189 Sogutlu, Sakarya, Turkey
A. Yildiz, Sakarya
VeriSeti Bilgi Teknolojileri Danışmanlık ve Ticaret, 44320 Adapazarı, Turkey

Data Set Information:
The dataset consists of session vectors belonging to 10,000 sessions. The dataset was formed as that each session corresponds to a different user. It is a time series data set, representing a specific campaign, spread over time, and split into groups.

Attribute Information:

<https://archive.ics.uci.edu/ml/datasets/Online+Shoppers+Purchasing+Intention+Dataset>

▼ Yomi Kastro

miuul

C.Okan Sakar ▼

simin
lduğu
N

Veri Setinin Yazarları

Online Shoppers Purchasing Intention Dataset, bir e-ticaret platformunun çevrimiçi ziyaretçilerinin satın alma niyetlerini incelemek için oluşturulmuş bir veri setidir. Veri seti, bir yıl boyunca toplanan kullanıcı oturumlarına ait çeşitli özellikleri içermektedir. Bu özellikler arasında ziyaretçilerin demografik bilgileri, site etkileşimleri, satın alma özellikleri, tarayıcı ve işletim sistemi bilgileri yer almaktadır. Veri seti, e-ticaret platformunun kullanıcı davranışlarını, pazarlama stratejilerini ve satın alma alışkanlıklarını anlamak için kullanılabilir. Bu veri seti üzerinde yapılan analizler, müşteri segmentasyonu, satın alma eğilimi tahmini ve çevrimiçi deneyim iyileştirmesi gibi konularda bilgi sağlayarak e-ticaret şirketlerine rekabet avantajı sunmaktadır.

miuul

VERİ
ANALİZİ

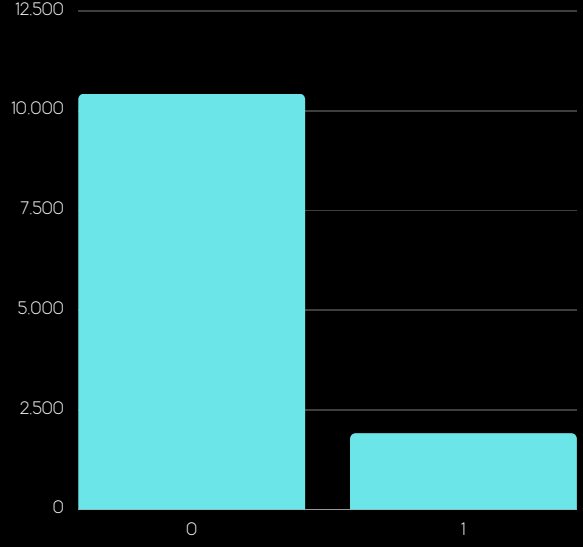


REVENUE

Hedef Değişken

Satış olup olmadığını ifade eden değişkendir

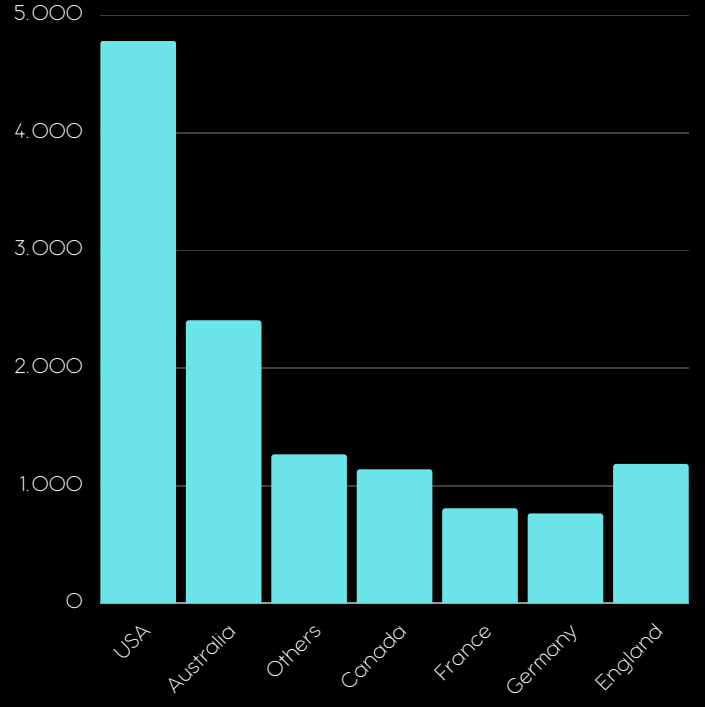
miuul



REGION

"Bölgeler"

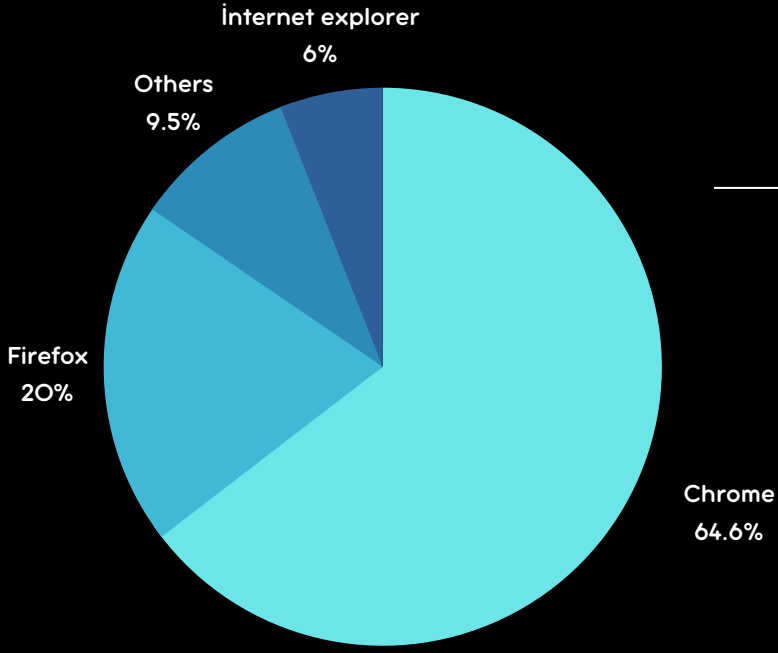
Site trafiğinin yoğunlukta olduğu ülkeler



BROWSER

"Tarayıcılar"

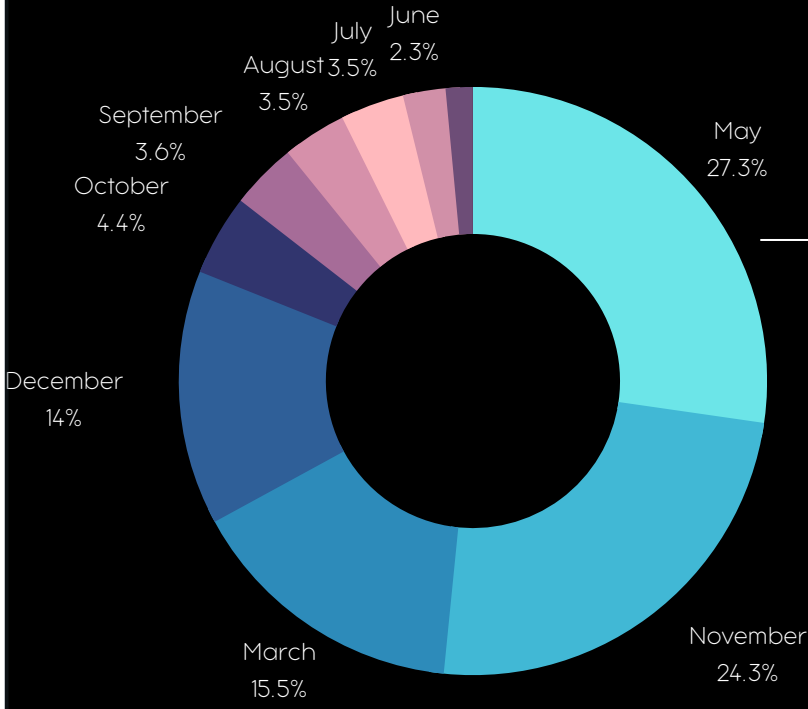
Müşterilerin kullandığı tarayıcı programları



MONTH

"Aylar"

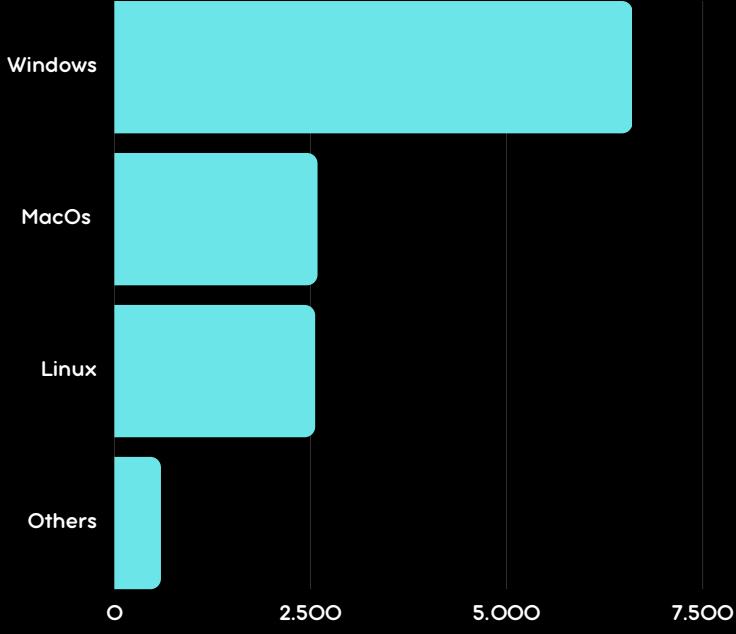
Aylara göre müşteri trafiği



OPERATING SYSTEMS

"İşletim Sistemleri"

Müşterilerin kullandığı işletim sistemleri



miuul

FEATURE ENGINEERING

- 'TotalTime' | Site üzerinde geçirilen toplam zaman
- Adm_time_per_page | Ana sayfa üzerinde geçirilen toplam zaman
- Pr_time_per_page | Kullanıcı sayfası üzerinde geçirilen toplam zaman





miuul

Machine Learning

Model Kurulumu ve Hiperparametrelerin Belirlenmesi

Ağaç Yöntemleri

Cross-Validate

Random-Forest

GridSearchCV

miuul

SONUÇLAR

HİPERPARAMETRE

SONUÇLAR

HİPERPARAMETRE

ÖNCESİ

RANDOM FOREST

- Accuracy: 0.8886
- Auc: 0.9051
- Recall: 0.4925
- Precision: 0.7297
- F1: 0.568

SONUÇLAR

HİPERPARAMETRE

ÖNCESİ

RANDOM FOREST

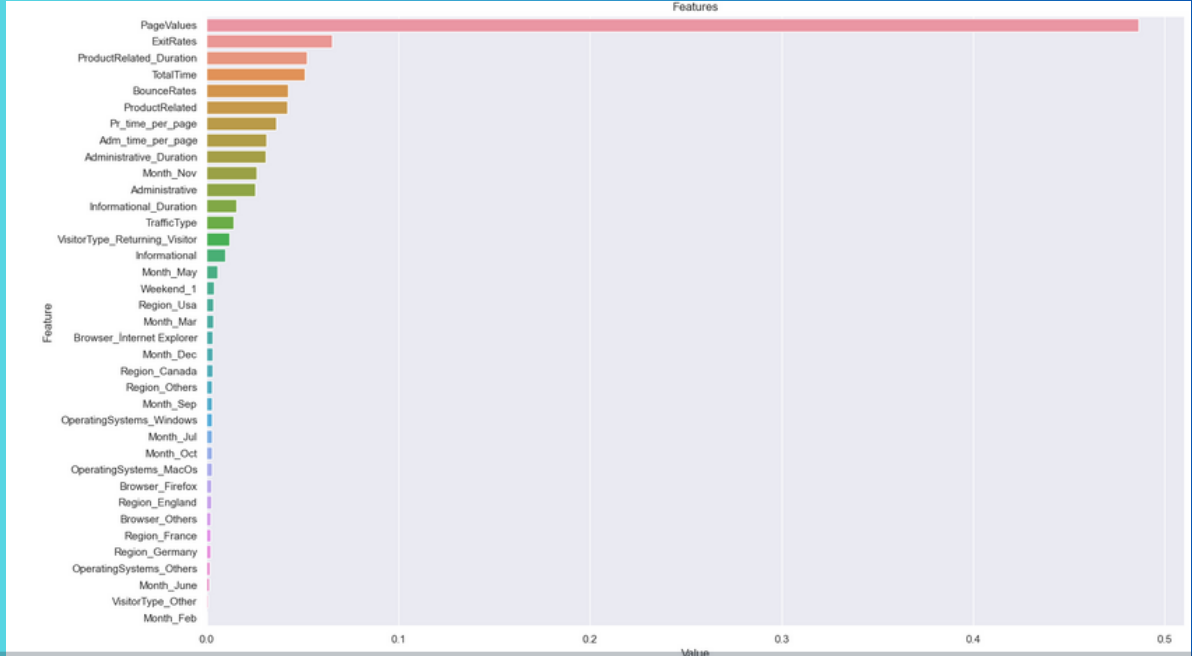
- Accuracy: 0.8886
- Auc: 0.9051
- Recall: 0.4925
- Precision: 0.7297
- F1: 0.568

SONRASI

RANDOM FOREST

- Accuracy: 0.8905
- Auc: 0.9104
- Recall: 5093
- Precision: 0.7245
- F1: 0.5805

SONUÇLAR



miuul

Teşekkür Ederiz