Model Report

Maqsad

Bu datasetda bizga moshina orindiqlar savdosi haqida umumiy malumot berilgan. Madsadi – shu mashina orindiqlarini AQSH da sotilsinmi yoki yoqmi aniqlash kerak.

Asosiy Natijalar

Bizga umumiy 10 ustunlik va ohirgi ustunini target variable qilb belgilab oldik hamda bizga 400 qatorlik dataset berilgan shular yordamida ishladik.

Biz umumiy 4 model bilan bu datasetni korib chiqdik ular: Random Forest, Decision Tree, Logistic Regression va KNN modellari. Har bir modelda har xil natijaga erishdik va natijaga erisha olmagan modellarimiz ham boldi.

Dataset Xusiyatlari Haqida Umumiy Malumot

Asosiy Xusiyatlar

Price: Moshina orindiqlar narxi

ShelveLocation: maxsulot joylashuvi

Urban: Maxsulot Shahar hududida sotilganmi yoki yoq

US: Maxsulot AQSH da sotilsinmi yoki yoq

Qoshimcha Xususiyatlari

Sales : Maxsulot savdosi

CompPrice : Raqobatchi maxsulotlar narxi

Income : Aholi daromadi

Advertising : Reklama uchun budjeti

Population : Aholi soni

Age : Aholi Yoshi

Education : Aholi talim darajasi

Malumotlar Tahlili:

Null qiymatlar- Datamizda Null qiymatlar bor bolgan ustunlar bor edi ular ShelveLocation hamda US edi. Biz ularni mode va mean valuelar bilan toldirib oldik.

Encoding – Biz Label Encoding yordamida ushbu object ustunlarni (Price, ShelveLoc, Education, Urban, US,) numeric valuega almashtirib oldik.

Scaling: Biz Standard Scaling orqali hamma xususiyatlarni bir xil qilb normallashtirib oldik

Target Variable : Target variable qilb “US” column ni tanlab oldik

Ishlatilgan Kutubxonalar Royxati

Decision Treeda	*Random Forest*	*Logistic Regression*	*KNN Model*
pandas	pandas	pandas	pandas
numpy	numpy	numpy	numpy
matplotlib.pyplot	matplotlib.pyplot	matplotlib.pyplot	matplotlib.pyplot
seaborn	seaborn	seaborn	seaborn
sklearn.model_selection	sklearn.model_selection	sklearn.model_selection	sklearn.model_selection
sklearn.tree	from sklearn.ensemble	from sklearn.ensemble	from sklearn.neighbors
sklearn.preprocessing	sklearn.preprocessing	fromsklearn.linear_model	sklearn.preprocessing
sklearn.metrics	sklearn.metrics	sklearn.metrics	sklearn.metrics

Modelni Rivojlantirish

Tanlangan modellar : Biz ushbu loyihamizni 2 da model orqali korib chqidik ular Random ***Forest, Decision Tree, Logistic Regression va KNN*** modellari.

Training the Model

Modelimizda biz datamizni training set a training set (70%) va testing set (30%) nisbatta split qilb modelimizni baholadik

Decision Treeda: modelni rivojlantirish uchun scallerdan foydalanib kordik

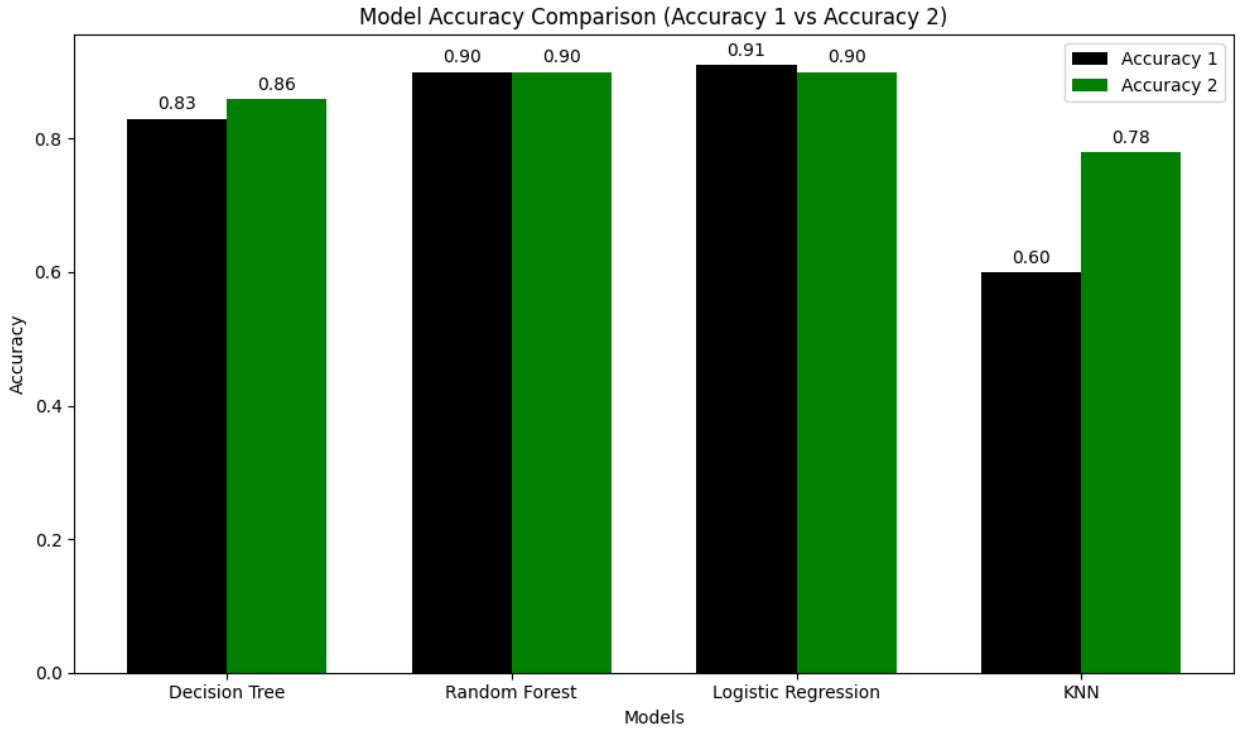
Random forest: Hyperparameter tuningdan foydalnaib rivojlantirib kordik (GridSearchCV)

Logistic Regression: Hyperparameter tuningdan foydalnaib rivojlantirib kordik (GridSearchCV)

KNN Model: k-fold K=10 cross-validation bolb oldik va foydalandik

Modelni Taqqoslash

Model Name	Natija (Accuracy 1)	Natija (Accuracy 2)
Decision Tree	0.83	0.86
Random Forest	0.90	0.90
Logistic Regression	0.91	0.90
KNN Model	0.60	0.78



Modelni Baholash

Modelni accuracyni baholashda biz calassification modelni accuracy = accuracy_score(y_test, y_pred) dan foydalandik

Modelni Endpointi

Modelni tamomladik va https://github.com/ShodiyAbdulloh/AI_ShodiyAbdulloh ga joylashtirdik

Xulosa

Logistic Regression va Random Forest modellarda biz deyarli hech qanday yangi natijaga erisha olmadik Hyperparameter tuninglarni ishlatib ham natija ololmadik

Decision Tree va K-Nearest Neighbors (KNN) modellari bizda yaxshi natija korsatdi. Birinchi accuracyni baholashda 0.83 va 0.60 natijalarga erishildi munosib ravishda va modelimizni rivojlantirgandan song esa munosib ravishda 0.86 va 0.78 natijalarga erishildi