



**YILDIZ TEKNİK ÜNİVERSİTESİ ELEKTRİK  
ELEKTRONİK FAKÜLTESİ  
BİLGİSAYAR MÜHENDİSLİĞİ**

Yapay Zeka Dersi

Ödev-2 Rapor

20011626 - Yusuf Yemliha ÇELİK

17011085 - Saliha KOÇYİĞİT

## Projenin Amacı:

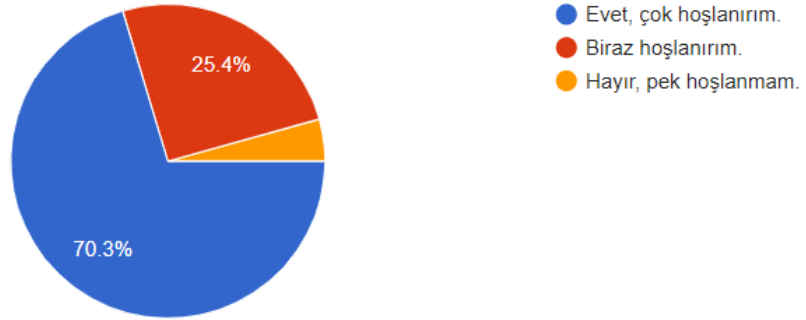
Bir veri kümesi oluşturarak farklı tahminleyicilerin hazırladığımız bu veri setindeki performansını analiz edip istatistiksel farkları t-test gibi farklı yöntemlerle kontrol ederek sonuçları tablo ve grafikler şeklinde sunmak.

## Veri Seti:

Bu veri seti, katılımcıların yeni deneyimlere olan ilgileri, sınırlarını zorlamaktan hoşlanma durumları, finansal risk alma eğilimleri, seyahat etme heyecanı, karar verme süreçleri, belirsizlikle başa çıkma yöntemleri ve diğer insanların risk alma davranışlarına yönelik düşünceleri gibi konulardaki tercihlerini içeren anket verilerinden oluşuyor. Son olarak katılımcının risk alma eğilimini değerlendirmesi için bir soru soruldu ve bu özellikler üzerinden model oluşturuldu. Verileri Google Forms üzerinden toplandı.

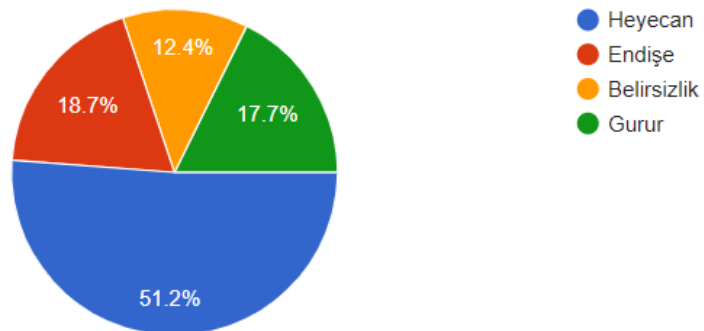
Yeni deneyimler yaşamaktan hoşlanır mısınız?

209 responses



Sınırlarınızı zorladığınızda, genellikle hissettiğiniz duygu hangisine daha yakındır?

209 responses



### Çıktılar:

- Bu çıktıları göre LogisticRegression, Precision ve Recall skorlarıyla diğer sınıflandırıcılara kıyasla daha iyi bir sonuç çıkarmıştır.
- DecisionTreeClassifier ve KNeighborsClassifier, diğer sınıflandırıcılara göre biraz daha düşük performans göstermiştir.
- SVC ve RandomForestClassifier benzer bir performans sergilemiştir.

Bu sonuçlara dayanarak **LogisticRegression** sınıflandırıcısının diğerlerine göre daha iyi bir performans sergilediği ve karar verme sürecinde daha iyi bir denge yakaladığı söylenebilir.

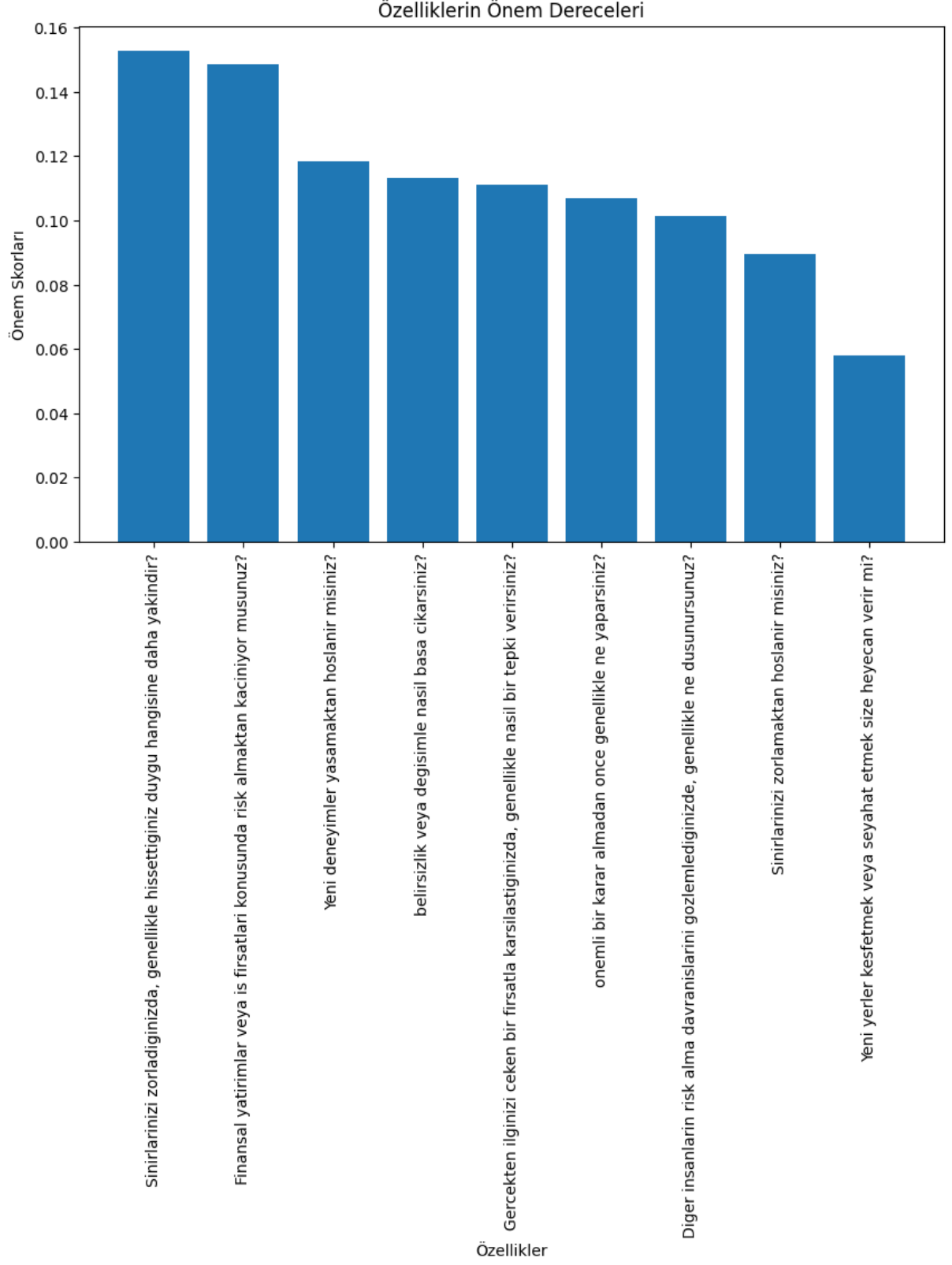
	Classifier	F1 Score	Precision Score	Accuracy Score	Recall Score
0	LogisticRegression	0.711329	0.725930	0.715000	0.715455
1	DecisionTreeClassifier	0.663768	0.666461	0.646905	0.655455
2	KNeighborsClassifier	0.631342	0.640277	0.636905	0.635000
3	SVC	0.698100	0.704899	0.700000	0.699091
4	RandomForestClassifier	0.686263	0.702407	0.713810	0.695909
5	AdaBoostClassifier	0.656666	0.669180	0.661429	0.662273

### T-Test Sonuçları:

T-test sonuçlarına göre, sınıflandırıcı çiftler arasında istatistiksel olarak anlamlı farklar tespit edilmemiştir. Ancak, “*LogisticRegression ve DecisionTreeClassifier*”, “*LogisticRegression ve KNeighborsClassifier*” arasında istatistiksel olarak anlamlı farklar olduğu görülmüştür. Bu sonuçlar, bu iki çiftin performansının diğerlerinden biraz daha farklı olduğunu gösterir.

T-Test Sonuçları:				
	Classifier 1	Classifier 2	t-Statistic	p-Value
0	LogisticRegression	DecisionTreeClassifier	1.709094	0.104618
1	LogisticRegression	KNeighborsClassifier	1.984072	0.062710
2	LogisticRegression	SVC	0.395218	0.697326
3	LogisticRegression	RandomForestClassifier	0.424197	0.676449
4	LogisticRegression	AdaBoostClassifier	1.529353	0.143561
5	DecisionTreeClassifier	KNeighborsClassifier	0.740568	0.468507
6	DecisionTreeClassifier	SVC	-1.224324	0.236613
7	DecisionTreeClassifier	RandomForestClassifier	-1.141612	0.268576
8	DecisionTreeClassifier	AdaBoostClassifier	0.310201	0.759971
9	KNeighborsClassifier	SVC	-1.759257	0.095520
10	KNeighborsClassifier	RandomForestClassifier	-2.007376	0.059964
11	KNeighborsClassifier	AdaBoostClassifier	-0.633698	0.534248
12	SVC	RandomForestClassifier	-1.073994	0.297013
13	SVC	AdaBoostClassifier	1.253722	0.225982
14	RandomForestClassifier	AdaBoostClassifier	1.390356	0.181375

## Özellikler ve model eğitimi için ağırlıkları



## Farklı Özellikler Arasındaki Korelasyon Matrisi:

Verilen korelasyon matrisi, değişkenler arasındaki ilişkileri değerlendirmek için kullanılmıştır. "Yeni yerler keşfetmek veya seyahat etmek size heyecan verir mi?" ve "Yeni deneyimler yaşamaktan hoşlanırmısınız?" özellikleri arasında 0.49 değerinde yüksek bir korelasyon tespit edilmiştir, yani bu iki özellik arasında güçlü bir ilişki olduğu söylenebilir.

