



# **DERİN ÖĞRENME İLE İNME TESPİTİ**

**Öğrencinin Adı - Soyadı :** Asuman Ulusoy - 21040101010

Efsa Sezer - 21040101008

Mert Can Sarıbiyık - 21040101025

(4908, 11)

|    | gender | age      | hypertension | heart_disease | ever_married | work_type | Residence_type | avg_glucose_level | bmi       | smoking_status | stroke |
|----|--------|----------|--------------|---------------|--------------|-----------|----------------|-------------------|-----------|----------------|--------|
| 0  | 1      | 1.069938 | 0            | 1             | 1            | 3         | 1              | 2.777797          | 0.981145  | 0              | 1      |
| 2  | 1      | 1.646336 | 0            | 1             | 1            | 3         | 0              | 0.014016          | 0.459086  | 1              | 1      |
| 3  | 0      | 0.271847 | 0            | 0             | 1            | 3         | 1              | 1.484266          | 0.701016  | 2              | 1      |
| 4  | 0      | 1.601998 | 1            | 0             | 1            | 4         | 0              | 1.549325          | -0.623231 | 1              | 1      |
| 5  | 1      | 1.690675 | 0            | 0             | 1            | 3         | 1              | 1.821493          | 0.013426  | 0              | 1      |
| 6  | 1      | 1.380306 | 1            | 1             | 1            | 3         | 0              | -0.792584         | -0.190304 | 1              | 1      |
| 7  | 0      | 1.158614 | 0            | 0             | 0            | 3         | 1              | -0.245546         | -0.776029 | 1              | 1      |
| 9  | 0      | 1.557660 | 0            | 0             | 1            | 3         | 1              | -1.051921         | -0.597765 | 3              | 1      |
| 10 | 0      | 1.690675 | 1            | 0             | 1            | 3         | 0              | -0.559812         | 0.102558  | 1              | 1      |
| 11 | 0      | 0.803907 | 0            | 1             | 1            | 1         | 0              | 0.341338          | 1.006611  | 2              | 1      |

# 1. Derin Öğrenme Modelinin Tespiti

## 1.1 Train – Test

Veri setimiz %70 train, %30 test olarak ayırdık.

```
from sklearn.model_selection import train_test_split,
cross_val_score,
GridSearchCV, RandomizedSearchCV, ShuffleSplit

x =
standardized_data.drop("stroke",
axis = 1) y =
```

**Çıktı :**

```
print(train_x.shape)
print(test_x.shape)
```

```
(3435, 10)
(1473, 10)
```

## 1.2 Uygun Modeli Bulma

LogisticRegression, DecisionTreeClassifier ve GradientBoostingRegressor modellerini araştırdık.

### 1.2.1 Logistic Regression

- **Ne İçin Kullanılır :** İkili sınıflandırma (binary classification) problemlerinde yaygın olarak kullanılır. Örneğin, bir e-postanın spam olup olmadığını belirlemek gibi.
- **Neden Kullanılır :** Basit ve hızlıdır, yüksek boyutlu veri setlerinde iyi performans gösterir. Ayrıca, sonuçları yorumlamak kolaydır.

### 1.2.2 Decision Tree Classifier

- **Ne İçin Kullanılır:** Sınıflandırma problemlerinde ve regresyon problemlerinde kullanılır. Karar ağaçları, bir dizi karar düğümü ve yaprak düğümü içerir. Her bir karar düğümü, bir özelliği test eder ve buna göre bir dalı izler.
- **Neden Kullanılır:** İyi tahmin performansı sağlar, aşırı uydurmaya karşı dirençlidir ve farklı türdeki veri setlerine uyarlanabilir. Ancak, parametre ayarı önemlidir.

### 1.2.3 Gradient Boosting Regressor

- **Ne İçin Kullanılır:** Regresyon problemlerinde kullanılır. Gradient Boosting, zayıf tahminçileri bir araya getirerek güçlü bir tahminci oluşturur.
- **Neden Kullanılır:** Basit bir modelleme yaklaşımı sunar, veri setlerini sınıflandırmak veya regresyon yapmak için yüksek performans sağlar. Ancak, aşırı uydurma (overfitting) eğilimindedir.

## 1.3 Modellerin Test Sonuçları

```
scores = gridSearch(x,  
y) scores
```

Çıktı :

|   | model                            | best_score | best_params                                       |
|---|----------------------------------|------------|---|
| 0 | L1 Reg                           | -0.001435  | {'alpha': 1}                                      |
| 1 | Decision Tree Classifier         | 0.923829   | {'splitter': 'best', 'criterion': 'entropy'}      |
| 2 | Gradient Boosting Classification | 0.023934   | {'n_estimators': 200, 'loss': 'squared_error',... |

## 2. Modelin Hazırlanması

```
Dtc_model = DecisionTreeClassifier(splitter = 'best', criterion =  
'entropy') Dtc_model.fit(train_x, train_y) # Eğitim  
Dtc_model.score(test_x, test_y) # Test
```

Çıktı :

0.9110658520027155

## 2.1 Confusion Matrix

Confusion Matrix, sınıflandırma problemlerinde modelin performansını değerlendirmek için kullanılan bir metriktir. Bu matris, gerçek sınıf ve tahmin edilen sınıf arasındaki ilişkiyi gösterir.

