

Schulungs- Cheat Sheet

1. Erstellung neuer Datenrahmen mit One hot-codierten Datensatz
`pd.get_dummies(„zuvor definierter Datensatz“, columns=[])`
2. Erstellung eines neuen Dataframes mit unabhängigen Variablen
`X = df.drop(" ",axis = 1)`
3. Erstellung der abhängigen Variablen
`Y = df[" "]`
4. Unterteilung der Daten in Test- & Trainingsdaten
Hier: 67% für Trainingsdaten & 33% für Testdaten
`xtrain, xtest, ytrain, ytest = train_test_split(X, Y, test_size=" ", random_state=42)`
5. Erstellung eines Sigmoid-Diagramms unter Verwendung der Daten der beliebigen Spalten (hier: „Fare“ & „Survived“)
`sns.Implot(x="Spalte 1", y="Spalte 2", data=data1,logistic=True, y_jitter=.03, height=15, aspect=1)`
6. Erstellung des Confusion Matrix
`confusion_matrix(ytest, pred)`
7. Accuracy (Berechnungsgenauigkeit)
`(tp + tn) / (tp + fp + tn + fn)`