from sklearn.datasets import make\_classification

import pandas as pd

# Create synthetic data for breast tumor classification

X, y = make\_classification(

n\_samples=1000, # Number of samples

n\_features=4, # Number of features

n\_informative=4, # All features are informative

n\_classes=2, # Binary classification: benign (0) vs malignant (1)

random\_state=42

)

# Assign meaningful names to the features

df = pd.DataFrame(X, columns=['Tumor\_Size', 'Tumor\_Texture', 'Tumor\_Area', 'Tumor\_Compactness'])

df['Target'] = y

# Display the first few rows of the dataset

print(df.head())

* **Tumor\_Size:** Size of the tumor in mm (continuous).
* **Tumor\_Texture:** Texture (smoothness or roughness of the tumor area, continuous).
* **Tumor\_Area:** The area of the tumor in mm² (continuous).
* **Tumor\_Compactness:** How compact the tumor is (calculated as Compactness=((Perimeter)2 / Area) − 1).

| **Tumor\_Size** | **Tumor\_Texture** | **Tumor\_Area** | **Tumor\_Compactness** | **Target** |
| --- | --- | --- | --- | --- |
| 1.600662 | 0.730988 | 1.370365 | 2.890064 | 0 (Benign) |
| 1.767906 | 2.450270 | 4.791903 | 0.000000 | 1 (Malignant) |
| 0.851814 | 0.000000 | 0.731659 | 3.306370 | 0 (Benign) |
| 2.244562 | 3.098886 | 2.217667 | 3.788212 | 1 (Malignant) |
| 4.149168 | 2.124404 | 0.000000 | 5.051919 | 1 (Malignant) |
| 2.463678 | 0.102427 | 0.718578 | 2.826321 | 1 (Malignant) |
| 0.000000 | 2.465510 | 2.930389 | 3.088816 | 0 (Benign) |
| 2.872880 | 1.603829 | 1.692938 | 2.826498 | 1 (Malignant) |
| 1.106042 | 3.931754 | 4.882092 | 1.526860 | 0 (Benign) |
| 1.119333 | 2.777160 | 3.469510 | 2.269244 | 0 (Benign) |