**Naive Bayes algorithm** for classification tasks like spam detection, sentiment analysis, or disease prediction.

**1. Play Tennis Dataset (Classic for Weather Prediction)**

**Use case:** Predict whether to play tennis or not based on weather conditions.

| **Outlook** | **Temperature** | **Humidity** | **Wind** | **PlayTennis** |
| --- | --- | --- | --- | --- |
| Sunny | Hot | High | Weak | No |
| Sunny | Hot | High | Strong | No |
| Overcast | Hot | High | Weak | Yes |
| Rain | Mild | High | Weak | Yes |
| Rain | Cool | Normal | Weak | Yes |
| Rain | Cool | Normal | Strong | No |
| Overcast | Cool | Normal | Strong | Yes |
| Sunny | Mild | High | Weak | No |
| Sunny | Cool | Normal | Weak | Yes |
| Rain | Mild | Normal | Weak | Yes |

Format: Categorical values  
Good for: Testing categorical Naive Bayes

**2. SMS Spam Collection Dataset**

**Use case:** Classify SMS messages as spam or not spam.  
Download: https://www.kaggle.com/datasets/uciml/sms-spam-collection-dataset

| **Label** | **Message** |
| --- | --- |
| ham | I'm going to attend the meeting today. |
| spam | WINNER!! You have won a $1000 gift card. Call now! |

Good for: Text classification using Multinomial Naive Bayes  
Preprocessing needed: Tokenization, vectorization (e.g., TF-IDF)

**3. Iris Dataset**

**Use case:** Predict flower species based on petal and sepal dimensions.  
Download: https://archive.ics.uci.edu/ml/datasets/iris

| **SepalLength** | **SepalWidth** | **PetalLength** | **PetalWidth** | **Species** |
| --- | --- | --- | --- | --- |
| 5.1 | 3.5 | 1.4 | 0.2 | setosa |
| 7.0 | 3.2 | 4.7 | 1.4 | versicolor |

Good for: Gaussian Naive Bayes  
Format: Numerical values

**4. Titanic Dataset**

**Use case:** Predict survival of passengers.  
Download: https://www.kaggle.com/competitions/titanic/data

| **PassengerId** | **Pclass** | **Sex** | **Age** | **Fare** | **Survived** |
| --- | --- | --- | --- | --- | --- |
| 1 | 3 | male | 22 | 7.25 | 0 |
| 2 | 1 | female | 38 | 71.28 | 1 |

Good for: Mixed features (categorical + numerical)

**5. Weather Dataset (Simplified)**

**Use case:** Predict rain based on weather features.

| **Outlook** | **Temp** | **Windy** | **Rain** |
| --- | --- | --- | --- |
| Sunny | Hot | False | No |
| Overcast | Cool | True | Yes |
| Rainy | Mild | False | Yes |