01 December 2022

UNIVERSITY OF GLOUCESTERSHIRE

# CT4031
# Maths for Data Science

## Week 6 – Practical

# Gaussian Classifier

```
from sklearn.datasets import load_iris
from sklearn.model_selection import train_test_split
from sklearn.naive_bayes import GaussianNB
X, y = load_iris(return_X_y=True)
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.5, random_state=0)
gnb = GaussianNB()
y_pred = gnb.fit(X_train, y_train).predict(X_test)
print("Number of mislabeled points")
print("Out of a total:", X_test.shape[0])
print("mislabeled: ", (y_test != y_pred).sum())
```

# Multinomial Classifier

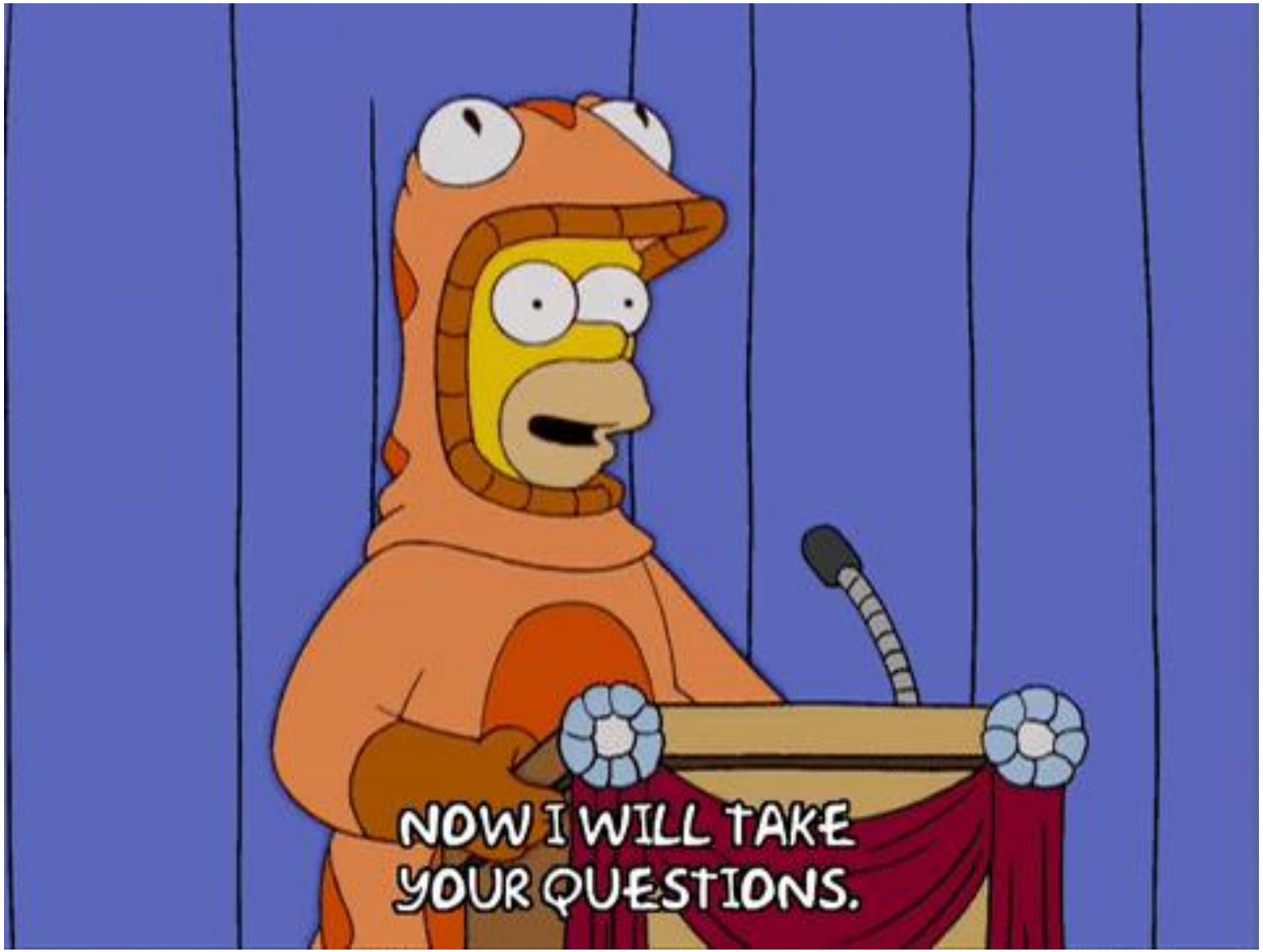
```
from sklearn.datasets import load_iris
from sklearn.model_selection import train_test_split
from sklearn.naive_bayes import MultinomialNB
X, y = load_iris(return_X_y=True)
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.5, random_state=0)
gnb = MultinomialNB()
y_pred = gnb.fit(X_train, y_train).predict(X_test)
print("Number of mislabeled points")
print("Out of a total:", X_test.shape[0])
print("mislabeled: ", (y_test != y_pred).sum())
```

# Bernoulli Classifier

```
from sklearn.datasets import load_iris
from sklearn.model_selection import train_test_split
from sklearn.naive_bayes import BernoulliNB
X, y = load_iris(return_X_y=True)
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.5, random_state=0)
gnb = BernoulliNB()
y_pred = gnb.fit(X_train, y_train).predict(X_test)
print("Number of mislabeled points")
print("Out of a total:", X_test.shape[0])
print("mislabeled: ", (y_test != y_pred).sum())
```

# Exercises

- Try using different datasets:
  - load_digits
  - load_wine


NOW I WILL TAKE
YOUR QUESTIONS.