

Informe del projecte de classificació de pingüins

URL del repositori GitHub:

https://github.com/FeryaelJustice/IEDIB_CursIABigData_TascaAvaluableCE_5073_3.1

Estructura del projecte

```
penguins_project/  
  README.md  
  environment.yml  
  models/  
    knn_model.pkl  
    logreg_model.pkl  
    svm_model.pkl  
    tree_model.pkl  
  client/  
    client.py  
  data/  
    penguins_size.csv  
  scripts/  
    train_knn.py  
    train_logreg.py  
    train_svm.py  
    train_tree.py  
    utils.py  
    __pycache__/  
      utils.cpython-311.pyc  
  app/  
    app.py
```

Contingut de environment.yml

```
name: penguins-env  
channels:  
  - defaults  
dependencies:  
  - python=3.9  
  - pandas  
  - seaborn  
  - scikit-learn  
  - flask  
  - requests  
  - numpy  
  - joblib  
  - pip
```

Execució del client

```
Testing model: logreg  
Sample 1 using local model (API unreachable)  
{  
  "class": 0,
```

```

    "species": "Adelie",
    "probabilities": [
      0.9969775597955256,
      0.0024222509447436352,
      0.0006001892597307583
    ]
  }
Sample 2 using local model (API unreachable)
{
  "class": 0,
  "species": "Adelie",
  "probabilities": [
    0.9934563693998775,
    0.005949595637878076,
    0.0005940349622443113
  ]
}
Testing model: svm
Sample 1 using local model (API unreachable)
{
  "class": 0,
  "species": "Adelie",
  "probabilities": [
    0.9918544451263396,
    0.003925854539258974,
    0.004219700334401474
  ]
}
Sample 2 using local model (API unreachable)
{
  "class": 0,
  "species": "Adelie",
  "probabilities": [
    0.9910567732514018,
    0.0039977529788696115,
    0.00494547376972864
  ]
}
Testing model: tree
Sample 1 using local model (API unreachable)
{
  "class": 0,
  "species": "Adelie",
  "probabilities": [
    1.0,
    0.0,
    0.0
  ]
}
Sample 2 using local model (API unreachable)
{
  "class": 0,
  "species": "Adelie",
  "probabilities": [

```

```

        1.0,
        0.0,
        0.0
    ]
}
Testing model: knn
Sample 1 using local model (API unreachable)
{
  "class": 0,
  "species": "Adelie",
  "probabilities": [
    1.0,
    0.0,
    0.0
  ]
}
Sample 2 using local model (API unreachable)
{
  "class": 0,
  "species": "Adelie",
  "probabilities": [
    1.0,
    0.0,
    0.0
  ]
}

```

Comentari dels resultats

Les quatre peticions corresponents a cada model retornen la mateixa classe (Adelie) amb probabilitats molt altes per a aquesta espècie. L'arbre i el KNN retornen una probabilitat de 1.0, mentre que la regressió logística i l'SVM assignen gairebé tota la massa de probabilitat a Adelie.