

Informe del projecte de classificació de pingüins

URL del repositori GitHub:

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

Estructura del projecte

The screenshot shows a code editor interface with the following details:

- File Explorer:** Shows the project structure:
 - PENGUINS_PROJECT
 - app
 - client
 - data
 - models
 - knn_model.pkl
 - logreg_model.pkl
 - svm_model.pkl
 - tree_model.pkl
 - scripts
 - _pycache_
 - train_knn.py
 - train_logreg.py
 - train_svm.py
 - train_tree.py
 - utils.py
 - .gitignore
 - environment.yml
 - penguins_report.pdf
 - README.md
- Code Editor:** The file `app.py` is open, showing Python code for a Flask application. It defines a route `/predict` that takes a model name as a parameter and returns a JSON response with predicted species and probabilities. The code uses scikit-learn models and a CSV dataset.
- Terminal:** Shows a git push command being run, compressing objects and writing them to a GitHub repository.
- Bottom Status Bar:** Displays the user's name (FeryaelJustice), session duration (48 minutes ago), and other system information.

```
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
```

environment.yml X

environment.yml

```
1 name: penguins-env
2 channels:
3   - defaults
4 dependencies:
5   - python=3.10
6   - pandas
7   - seaborn
8   - scikit-learn
9   - flask
10  - requests
11  - numpy
12  - joblib
13  - pip
```

```
nano9 ~\OneDrive\Escritorio\Stuff\Professional\Academical\Specialization\AIBigData\AIProgramming\Unit3\penguins_project > python .\client\client.py

Testing model: logreg
Sample 1 response: {
  "class": 0,
  "probabilities": [
    0.9969775597955256,
    0.002422250944743631,
    0.0006001892597307573
  ],
  "species": "Adelie"
}
Sample 2 response: {
  "class": 0,
  "probabilities": [
    0.9934563693998775,
    0.005949595637878066,
    0.0005940349622443107
  ],
  "species": "Adelie"
}

Testing model: svm
Sample 1 response: {
  "class": 0,
  "probabilities": [
    0.992117323751276,
    0.0037582054368801117,
    0.004124470811843848
  ],
  "species": "Adelie"
}
Sample 2 response: {
  "class": 0,
  "probabilities": [
    0.9913421710064303,
    0.0038336365644800283,
    0.004824192429089887
  ],
  "species": "Adelie"
}

Testing model: tree
Sample 1 response: {
  "class": 0,
  "probabilities": [
    1.0,
    0.0,
    0.0
  ],
  "species": "Adelie"
}
Sample 2 response: {
  "class": 0,
  "probabilities": [
    1.0,
    0.0,
    0.0
  ],
  "species": "Adelie"
}

Testing model: knn
Sample 1 response: {
  "class": 0,
  "probabilities": [
    1.0,
    0.0,
    0.0
  ],
  "species": "Adelie"
}
Sample 2 response: {
  "class": 0,
  "probabilities": [
    1.0,
    0.0,
    0.0
  ],
  "species": "Adelie"
}
```

```
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.