

Fairness Analyze

Endpoint: `/api/v1/fairness/Analyze`

Payload Details:

biasType: Provide bias type based on your requirement PRETRAIN/POSTTRAIN.

methodType: Provided method type for metric score like disparate impact or ALL will return available metric scores.

taskType: As of now we have only CLASSIFICATION.

Label: Mention the target column name to predict.

predLabel: Add prediction label. Default is “labels_pred”. This is required for POSTTRAIN.

FavourableOutcome: Mention favorable outcome for predict column.

ProtectedAttribute: Mention the protected attribute column name.

Privileged: Mention the privileged value for protected attribute. If multiple Privileged groups are there, entered in this format [priv_1,priv_2],[priv_3,priv_4]

File: Upload the dataset.

Infosys Responsible AI - Analyze UploadFile

POST /api/v1/fairness/Analyze Analyze Uploadfile

Parameters

No parameters

Request body required

multipart/form-data

biasType * required string PRETRAIN

methodType * required string ALL

taskType * required string CLASSIFICATION

Label * required string income-per-year

predLabel * required string income-per-year

FavourableOutcome * required string >50K

ProtectedAttribute * required string race

Privileged * required string White

file * required string(Binary) Choose File adult.csv

Execute

Response:

Code	Details
200	<div><div>Response body</div><div><pre>{ "biasResults": [{ "biasDetected": true, "protectedAttribute": [{ "name": "race", "privileged": ["White"], "unprivileged": ["Black", "Asian-Pac-Islander", "Other", "Amer-Indian-Eskimo"] }], "metrics": [{ "name": "STATISTICAL PARITY DIFFERENCE", "description": "The difference in the rate of favorable outcomes received by unprivileged group to the privileged group. Ideal value for this is 0, which means there is no biasness present. Negative value for this means that the data is biased towards the privileged group and positive values means, it is biased towards the unprivileged group.", "value": "-0.1" }, { "name": "DISPARATE IMPACT", </pre></div><div><div>Download</div></div></div> <div><div>Response headers</div><div><pre>access-control-allow-origin: http://10.66.155.13:30005 content-length: 1395 content-type: application/json date: Sat, 13 Jul 2024 12:12:02 GMT server: uvicorn vary: Origin</pre></div></div>

For Post Train

POST

/api/v1/fairness/Analyse

Analyse Uploadfile

Parameters

No parameters

Request body

required

multipart/form-data

biasType

string

methodType

string

taskType

string

Label

string

predLabel

string

FavourableOutcome

string

ProtectedAttribute

string

Privileged

string

file

string(\$binary)

POSTTRAIN

ALL

CLASSIFICATION

income-per-year

labels_pred

>50K

race

White

Choose file

adult.csv

Execute

Response:


Code

Details

200

Response body

```
{
  "biasResults": [
    {
      "biasDetected": true,
      "protectedAttribute": [
        {
          "name": "race",
          "privileged": [
            "white"
          ],
          "unprivileged": [
            "black",
            "Asian-Pac-Islander",
            "other",
            "Amer-Indian-Eskimo"
          ]
        }
      ],
      "metrics": [
        {
          "name": "Statistical parity",
          "description": "This function computes the statistical parity (difference of success rates) between group_unprivileged and group_privileged. A value of 0 is desired. Negative values are unfair towards group_unprivileged. Positive values are unfair towards group_privileged. The range (-0.1,0.1) is considered acceptable.",
          "value": "-0.05"
        },
        {
          "name": "Disparate Impact",
```

 [Download](#)

Response headers

```
access-control-allow-origin: http://10.66.155.13:30005
content-length: 2002
content-type: application/json
date: Sat,13 Jul 2024 12:19:03 GMT
server: uvicorn
vary: Origin
```

Responses

In-Processing Analyze:

Endpoint: `api/v1/fairness/inprocessing/exponentiated_gradient_reduction`

Step1: Please go to the "exponentiated_gradient_reduction" API at the URL mentioned above.

POST `/api/v1/fairness/inprocessing/exponentiated_gradient_reduction` Inprocessing Exponentiated Gradient Reduction

trainingDataset * required
string(\$binary)

Choose File No file chosen

testingDataset * required
string(\$binary)

Choose File No file chosen

label * required
string

income-per-year

favourableOutcome * required
string

1

sensitiveFeatures * required
string

race

Label: Mention the target column name to predict.

FavourableOutcome: Mention favourable outcome for predict column.

ProtectedAttribute: Mention the protected attribute column name.

Provide datasets required to execute.

Code	Details
200	<p>Response body</p> <pre>{ "modelName": "aware_model_06252024081147.joblib", "metrics": { "demographic_parity_difference": 0.205579123604275, "equalized_odds_difference": 0.4396423248882265, "true_positive_rate": 0.6550365785030952, "true_negative_rate": 0.930990990990991, "false_positive_rate": 0.06900900900900901, "false_negative_rate": 0.3449634214969049, "accuracy_score": 0.8640644192711887 } }</pre> <p>Response headers</p>

You will receive the return of metric scores, return optimized model name as a response.

Step2: Please go to the "getModel/{filename}" API at the URL mentioned above.

GET	/api/v1/fairness/inprocessing/getModel/{filename} Inprocessing Get Model
-----	--

Provide the filename to download the model.

Code	Details
200	<p>Response body</p> <p>Download file</p> <p>Response headers</p>

Click on download file to save into local.

Individual metrics:

Step1:

Endpoint: `api/v1/fairness/individual/bias/getlabels`

Please go to the "getlabels" API at the URL mentioned above.

POST `/api/v1/fairness/individual/bias/getlabels` Get Labels

No parameters

Request body required

fileid * required
number

Provide the file id to return attributes for dataset.

request URL

<https://api-aicloud.ad.infosys.com/api/v1/fairness/individual/bias/getlabels>

Server response

Code

Details

200

Response body

```
{
  "FileName": "adultTrainData.csv",
  "UploadedFileType": "text/csv",
  "AttributesInTheDataset": {
    "FeatureList ": [
      "age",
      "workclass",
      "fnlwgt",
      "education",
      "education-num",
      "marital-status",
      "occupation",
      "relationship",
      "race",
      "sex",
      "capital-gain",
      "capital-loss",
      "hours-per-week",
      "native-country",
      "income-per-year"
    ]
  },
  "recordId": "667913342fb10a116b2d7727"
}
```

Step2:

Please go to the "getscore" API at the URL mentioned above.

POST

[/api/v1/fairness/individual/bias/getscore](#) Get Individualscore

Request body **required**


```
{
  "labels": [
    "income-per-year"
  ],
  "recordId": "667913342fb10a116b2d7727|"
}
```

Provide the label and recordId to execute.

Request URL

```
https://api-aicloud.ad.infosys.com/api/v1/fairness/individual/bias/getscore
```

Server response

Code	Details
200	<p>Response body</p> <pre>[{ "income-per-year": { "name": "CONSISTENCY", "description": "Individual fairness metric that measures how similar the labels are for similar instances. Score ranges from [0,1], where 1 indicates consistent 1 labels for similar instances.", "value": 0.78 } }]</pre> <div> Download</div>

You will receive the metrics score for provided dataset.

Mitigate Dataset:

Step-1

Endpoint: `api/v1/fairness/pretrain/mitigation/getDataset`

POST `/api/v1/fairness/pretrain/mitigation/getDataset` Get Data

Request body required

MitigationType <small>* required</small> string	PREPROCESSING
MitigationTechnique <small>* required</small> string	REWEIGHING
taskType <small>* required</small> string	CLASSIFICATION
fileId <small>* required</small> number	1712558644.746234

MitigationType: Mention the mitigationType Preprocessing.

MitigationTechnique: Mention the mitigationTechnique to mitigate.

taskType: As of now we have only CLASSIFICATION.

fileId: provide file id from database.

After entering all the necessary data, proceed by clicking on "execute."

You will receive the return of all attributes for provided dataset as a response.

Request URL

<https://api-aicloud.ad.infosys.com/api/v1/fairness/pretrain/mitigation/getDataset>

Server response

Code

Details

200

Response body

```
{
  "mitigationType": "PREPROCESSING",
  "mitigationTechnique": "REWEIGHING",
  "trainFileName": "adultTrainData.csv",
  "UploadedFileType": "text/csv",
  "AttributesInTheDataset": {
    "FeatureList ": [
      "age",
      "workclass",
      "fnlwgt",
      "education",
      "education-num",
      "marital-status",
      "occupation",
      "relationship",
      "race",
      "sex",
      "capital-gain",
      "capital-loss",
      "hours-per-week",
      "native-country",
      "income-per-year"
    ],
    "CategoricalAttributesList": [
      "workclass",
      "education",
      "marital-status",
      "occupation",

```

Response headers

Step-2:

Please go to the "pretrain/mitigation/getAttributes" API at the URL mentioned above

POST

</api/v1/fairness/pretrain/mitigation/getAttributes> Get Attributes

Request body required

Label * required

string

income-per-year

FavourableOutcome * required

string

>50K

ProtectedAttribute * required

string

race

Priviledged * required

string

White

recordId * required

string

66790c542fb10a116b2d7676

Provide necessary data to execute.

Label: Mention the target column name to predict.

FavourableOutcome: Mention favourable outcome for predict column.

ProtectedAttribute: Mention the protected attribute column name.

Privileged: Mention the privileged value for protected attribute.

recordId: Provide the record id which is returned in getDataset Api along with attributes.

You will receive the metrics score for provided dataset and mitigated filename as a response.

Request URL

`https://api-aicloud.ad.infosys.com/api/v1/fairness/pretrain/mitigation/getAttributes`

Server response

Code	Details
200	<p>Response body</p> <pre>{ "biasResults": [{ "biasDetected": true, "protectedAttribute": [{ "name": "race", "privileged": ["White"], "unprivileged": ["Black", "Asian-Pac-Islander", "Other", "Amer-Indian-Eskimo"] }], "metrics": [{ "name": "STATISTICAL PARITY-DIFFERENCE", "description": "The difference in the rate of favorable outcomes received by unprivileged group to the privileged group. Ideal value for this is 0, which means there is no biasness present. Negative value for this means that the data is biased towards the privileged group and positive values means, it is biased towards the unprivileged group.", "value": "-0.1" }, { "name": "DISPARATE-IMPACT", </pre> <p>Response headers</p>

Step3:

GET `/api/v1/fairness/mitigation/getMitigatedData/{fileName}` Process

Use above endpoint to download mitigate dataset by provided filename.

Server response

Code	Details
200	<p>Response body</p> <p>Download file</p> <p>Response headers</p> <pre>connection: keep-alive content-disposition: attachment; filename=PMitigated_adultTrainData.csv_06242024062721csv content-length: 21822148 date: Mon, 24 Jun 2024 06:27:42 GMT etag: "59df96c5d8b558a1567eea47a0e6df86" last-modified: Mon, 24 Jun 2024 06:27:22 GMT server: nginx/1.19.5 x-envoy-upstream-service-time: 2</pre>

Click on download file to download the into local.