Generate | Using ... | create a dataframe with 2 columns and 10 rows | Close

```
!pip install tensorflow_decision_forests wurlitzer
```

```
Collecting tensorflow_decision_forests
  Downloading tensorflow_decision_forests-1.8.1-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (15.3 MB)
     ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 15.3/15.3 MB 23.3 MB/s eta 0:00:00
Collecting wurlitzer
  Downloading wurlitzer-3.0.3-py3-none-any.whl (7.3 kB)
Requirement already satisfied: numpy in /usr/local/lib/python3.10/dist-packages (from tensorflow_decision_forests) (1.23.5)
Requirement already satisfied: pandas in /usr/local/lib/python3.10/dist-packages (from tensorflow_decision_forests) (1.5.3)
Requirement already satisfied: tensorflow~=2.15.0 in /usr/local/lib/python3.10/dist-packages (from tensorflow_decision_forests) (2.15.0)
Requirement already satisfied: six in /usr/local/lib/python3.10/dist-packages (from tensorflow_decision_forests) (1.16.0)
Requirement already satisfied: absl-py in /usr/local/lib/python3.10/dist-packages (from tensorflow_decision_forests) (1.4.0)
Requirement already satisfied: wheel in /usr/local/lib/python3.10/dist-packages (from tensorflow_decision_forests) (0.42.0)
Requirement already satisfied: astunparse>=1.6.0 in /usr/local/lib/python3.10/dist-packages (from tensorflow~=2.15.0->tensorflow_decision_forests) (1.6
Requirement already satisfied: flatbuffers>=23.5.26 in /usr/local/lib/python3.10/dist-packages (from tensorflow~=2.15.0->tensorflow_decision_forests) (
Requirement already satisfied: gast!=0.5.0,!=0.5.1,!=0.5.2,>=0.2.1 in /usr/local/lib/python3.10/dist-packages (from tensorflow~=2.15.0->tensorflow_deci
Requirement already satisfied: google-pasta>=0.1.1 in /usr/local/lib/python3.10/dist-packages (from tensorflow~=2.15.0->tensorflow_decision_forests) (0
Requirement already satisfied: h5py>=2.9.0 in /usr/local/lib/python3.10/dist-packages (from tensorflow~=2.15.0->tensorflow_decision_forests) (3.9.0)
Requirement already satisfied: libclang>=13.0.0 in /usr/local/lib/python3.10/dist-packages (from tensorflow~=2.15.0->tensorflow_decision_forests) (16.0
Requirement already satisfied: ml-dtypes~=0.2.0 in /usr/local/lib/python3.10/dist-packages (from tensorflow~=2.15.0->tensorflow_decision_forests) (0.2.(
Requirement already satisfied: opt-einsum>=2.3.2 in /usr/local/lib/python3.10/dist-packages (from tensorflow~=2.15.0->tensorflow_decision_forests) (3.3
Requirement already satisfied: packaging in /usr/local/lib/python3.10/dist-packages (from tensorflow~=2.15.0->tensorflow_decision_forests) (23.2)
Requirement already satisfied: protobuf!=4.21.0,!=4.21.1,!=4.21.2,!=4.21.3,!=4.21.4,!=4.21.5,<5.0.0dev,>=3.20.3 in /usr/local/lib/python3.10/dist-packa
Requirement already satisfied: setuptools in /usr/local/lib/python3.10/dist-packages (from tensorflow~=2.15.0->tensorflow_decision_forests) (67.7.2)
Requirement already satisfied: termcolor>=1.1.0 in /usr/local/lib/python3.10/dist-packages (from tensorflow~=2.15.0->tensorflow_decision_forests) (2.4.(
Requirement already satisfied: typing-extensions>=3.6.6 in /usr/local/lib/python3.10/dist-packages (from tensorflow~=2.15.0->tensorflow_decision_forest
Requirement already satisfied: wrapt<1.15,>=1.11.0 in /usr/local/lib/python3.10/dist-packages (from tensorflow~=2.15.0->tensorflow_decision_forests) (1
Requirement already satisfied: tensorflow-io-gcs-filesystem>=0.23.1 in /usr/local/lib/python3.10/dist-packages (from tensorflow~=2.15.0->tensorflow_dec
Requirement already satisfied: grpcio<2.0,>=1.24.3 in /usr/local/lib/python3.10/dist-packages (from tensorflow~=2.15.0->tensorflow_decision_forests) (1
Requirement already satisfied: tensorboard<2.16,>=2.15 in /usr/local/lib/python3.10/dist-packages (from tensorflow~=2.15.0->tensorflow_decision_forests
Requirement already satisfied: tensorflow-estimator<2.16,>=2.15.0 in /usr/local/lib/python3.10/dist-packages (from tensorflow~=2.15.0->tensorflow_decis
Requirement already satisfied: keras<2.16,>=2.15.0 in /usr/local/lib/python3.10/dist-packages (from tensorflow~=2.15.0->tensorflow_decision_forests) (2
Requirement already satisfied: python-dateutil>=2.8.1 in /usr/local/lib/python3.10/dist-packages (from pandas->tensorflow_decision_forests) (2.8.2)
Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.10/dist-packages (from pandas->tensorflow_decision_forests) (2023.3.post1)
Requirement already satisfied: google-auth<3,>=1.6.3 in /usr/local/lib/python3.10/dist-packages (from tensorboard<2.16,>=2.15->tensorflow~=2.15.0->tens
Requirement already satisfied: google-auth-oauthlib<2,>=0.5 in /usr/local/lib/python3.10/dist-packages (from tensorboard<2.16,>=2.15->tensorflow~=2.15.(
Requirement already satisfied: markdown>=2.6.8 in /usr/local/lib/python3.10/dist-packages (from tensorboard<2.16,>=2.15->tensorflow~=2.15.0->tensorflow_
Requirement already satisfied: requests<3,>=2.21.0 in /usr/local/lib/python3.10/dist-packages (from tensorboard<2.16,>=2.15->tensorflow~=2.15.0->tensor
Requirement already satisfied: tensorboard-data-server<0.8.0,>=0.7.0 in /usr/local/lib/python3.10/dist-packages (from tensorboard<2.16,>=2.15->tensorfl
Requirement already satisfied: werkzeug>=1.0.1 in /usr/local/lib/python3.10/dist-packages (from tensorboard<2.16,>=2.15->tensorflow~=2.15.0->tensorflow_
Requirement already satisfied: cachetools<6.0,>=2.0.0 in /usr/local/lib/python3.10/dist-packages (from google-auth<3,>=1.6.3->tensorboard<2.16,>=2.15->
Requirement already satisfied: pyasn1-modules>=0.2.1 in /usr/local/lib/python3.10/dist-packages (from google-auth<3,>=1.6.3->tensorboard<2.16,>=2.15->t
Requirement already satisfied: rsa<5,>=3.1.4 in /usr/local/lib/python3.10/dist-packages (from google-auth<3,>=1.6.3->tensorboard<2.16,>=2.15->tensorflo
Requirement already satisfied: requests-oauthlib>=0.7.0 in /usr/local/lib/python3.10/dist-packages (from google-auth-oauthlib<2,>=0.5->tensorboard<2.16
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.10/dist-packages (from requests<3,>=2.21.0->tensorboard<2.16,>=2.15->
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.10/dist-packages (from requests<3,>=2.21.0->tensorboard<2.16,>=2.15->tensorflow~=
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.10/dist-packages (from requests<3,>=2.21.0->tensorboard<2.16,>=2.15->tensor
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.10/dist-packages (from requests<3,>=2.21.0->tensorboard<2.16,>=2.15->tensor
Requirement already satisfied: MarkupSafe>=2.1.1 in /usr/local/lib/python3.10/dist-packages (from werkzeug>=1.0.1->tensorboard<2.16,>=2.15->tensorflow~
Requirement already satisfied: pyasn1<0.6.0,>=0.4.6 in /usr/local/lib/python3.10/dist-packages (from pyasn1-modules>=0.2.1->google-auth<3,>=1.6.3->tens
Requirement already satisfied: oauthlib>=3.0.0 in /usr/local/lib/python3.10/dist-packages (from requests-oauthlib>=0.7.0->google-auth-oauthlib<2,>=0.5-
Installing collected packages: wurlitzer, tensorflow_decision_forests
Successfully installed tensorflow_decision_forests-1.8.1 wurlitzer-3.0.3
```

```
!pip install rasterio
```

```
Collecting rasterio
  Downloading rasterio-1.3.9-cp310-cp310-manylinux2014_x86_64.whl (20.6 MB)
     ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 20.6/20.6 MB 52.2 MB/s eta 0:00:00
Collecting affine (from rasterio)
  Downloading affine-2.4.0-py3-none-any.whl (15 kB)
Requirement already satisfied: attrs in /usr/local/lib/python3.10/dist-packages (from rasterio) (23.2.0)
Requirement already satisfied: certifi in /usr/local/lib/python3.10/dist-packages (from rasterio) (2023.11.17)
Requirement already satisfied: click>=4.0 in /usr/local/lib/python3.10/dist-packages (from rasterio) (8.1.7)
Requirement already satisfied: cligj>=0.5 in /usr/local/lib/python3.10/dist-packages (from rasterio) (0.7.2)
Requirement already satisfied: numpy in /usr/local/lib/python3.10/dist-packages (from rasterio) (1.23.5)
Collecting snuggs>=1.4.1 (from rasterio)
  Downloading snuggs-1.4.7-py3-none-any.whl (5.4 kB)
Requirement already satisfied: click-plugins in /usr/local/lib/python3.10/dist-packages (from rasterio) (1.1.1)
Requirement already satisfied: setuptools in /usr/local/lib/python3.10/dist-packages (from rasterio) (67.7.2)
Requirement already satisfied: pyparsing>=2.1.6 in /usr/local/lib/python3.10/dist-packages (from snuggs>=1.4.1->rasterio) (3.1.1)
Installing collected packages: snuggs, affine, rasterio
Successfully installed affine-2.4.0 rasterio-1.3.9 snuggs-1.4.7
```

```
from google.colab import drive
drive.mount('/content/drive')
```

```
Mounted at /content/drive
```

```
import tensorflow_decision_forests as tfdf

import os
import numpy as np
import pandas as pd
import tensorflow as tf
import math

from IPython.core.magic import register_line_magic
from IPython.display import Javascript
from IPython.display import display as ipy_display

import rasterio
import pandas as pd
from sklearn.model_selection import train_test_split
```

```
# Some of the model training logs can cover the full
# screen if not compressed to a smaller viewport.
# This magic allows setting a max height for a cell.
@register_line_magic
def set_cell_height(size):
  ipy_display(
      Javascript("google.colab.output.setIframeHeight(0, true, {maxHeight: " +
                 str(size) + "})"))


# Check the version of TensorFlow Decision Forests
print("Found TensorFlow Decision Forests v" + tfdf.__version__)
```

```
Found TensorFlow Decision Forests v1.8.1
```

```
grupo_path = '/content/drive/MyDrive/data/output_grupo_raster.tif'
binary_path = '/content/drive/MyDrive/data/binary_deforestation_raster.tif'
```

```
# Load the raster data
with rasterio.open(grupo_path) as src:
    feature_raster = src.read(1)

with rasterio.open(binary_path) as src:
    label_raster = src.read(1)
```

```
# Flatten the raster arrays and create a DataFrame
pixels = {'feature': feature_raster.flatten(),
          'label': label_raster.flatten()}
df = pd.DataFrame(pixels)

# Remove nodata pixels (assuming nodata is -1 for label)
df = df[df['label'] != -1]
```

```
# Split the data into training and testing sets
train_df, test_df = train_test_split(df, test_size=0.3)
```



```
del df
del pixels
```

```
# Convert DataFrame to TensorFlow dataset
train_ds = tfdf.keras.pd_dataframe_to_tf_dataset(train_df, label='label')
test_ds = tfdf.keras.pd_dataframe_to_tf_dataset(test_df, label='label')
```

```
# Train a Random Forest model
model = tfdf.keras.RandomForestModel()
model.fit(train_ds)
```

```
Use /tmp/tmp8qr8ocr2 as temporary training directory
Reading training dataset...
Training dataset read in 0:01:27.313504. Found 59064655 examples.
Training model...
Model trained in 0:24:07.327010
Compiling model...
Model compiled.
<keras.src.callbacks.History at 0x78da5c7146d0>
```

```
# Evaluate the model
evaluation = model.evaluate(test_ds)

print(evaluation)
```

```
25314/25314 [==============================] - 85s 3ms/step - loss: 0.0000e+00
0.0
```

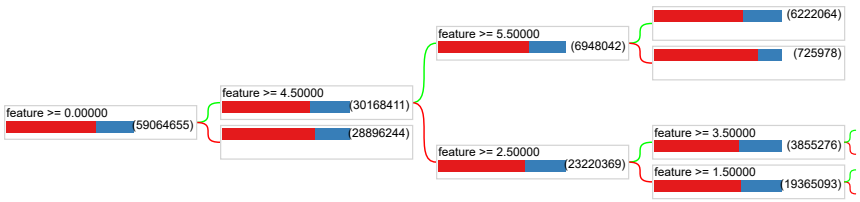
```
model.compile(metrics=["accuracy"])
evaluation = model.evaluate(test_ds, return_dict=True)
print()

for name, value in evaluation.items():
  print(f"{name}: {value:.4f}")
```

```
25314/25314 [==============================] - 87s 3ms/step - loss: 0.0000e+00 - accuracy: 0.7076

loss: 0.0000
accuracy: 0.7076
```

```
tfdf.model_plotter.plot_model_in_colab(model, tree_idx=0, max_depth=3)
```




```
%set_cell_height 300
model.summary()
```

```
trees: 263, Out-of-bag evaluation: accuracy:0.707625 logloss:10.5383
trees: 265, Out-of-bag evaluation: accuracy:0.707625 logloss:10.5383
trees: 268, Out-of-bag evaluation: accuracy:0.707625 logloss:10.5383
trees: 270, Out-of-bag evaluation: accuracy:0.707625 logloss:10.5383
trees: 272, Out-of-bag evaluation: accuracy:0.707625 logloss:10.5383
trees: 274, Out-of-bag evaluation: accuracy:0.707625 logloss:10.5383
trees: 277, Out-of-bag evaluation: accuracy:0.707625 logloss:10.5383
trees: 279, Out-of-bag evaluation: accuracy:0.707625 logloss:10.5383
trees: 282, Out-of-bag evaluation: accuracy:0.707625 logloss:10.5383
trees: 284, Out-of-bag evaluation: accuracy:0.707625 logloss:10.5383
trees: 286, Out-of-bag evaluation: accuracy:0.707625 logloss:10.5383
trees: 289, Out-of-bag evaluation: accuracy:0.707625 logloss:10.5383
trees: 291, Out-of-bag evaluation: accuracy:0.707625 logloss:10.5383
trees: 294, Out-of-bag evaluation: accuracy:0.707625 logloss:10.5383
trees: 297, Out-of-bag evaluation: accuracy:0.707625 logloss:10.5383
trees: 300, Out-of-bag evaluation: accuracy:0.707625 logloss:10.5383
```

Accuracy: The OOB accuracy of 0.707625 suggests that about 70.76% of the OOB samples were correctly classified by the model. Logloss: The logarithmic loss (logloss) of 10.5383 is a measure of error and is more sensitive to classifiers that are confident about an incorrect classification. OOB evaluation is a method of measuring the prediction error of random forest models. For each tree, it uses only the data that was not included in the bootstrap sample (the "out-of-bag" data) to evaluate the model's performance. This provides a good estimate of how well the model might perform on unseen data.

```
# The input features
model.make_inspector().features()
```

```
     ["feature" (1; #1)]
```

```
# The feature importances
model.make_inspector().variable_importances()
```

```
     {'SUM_SCORE': [("feature" (1; #1), 33616550.91884154)],
      'INV_MEAN_MIN_DEPTH': [("feature" (1; #1), 1.0)],
      'NUM_NODES': [("feature" (1; #1), 1770.0)],
      'NUM_AS_ROOT': [("feature" (1; #1), 300.0)]}
```

```
model.make_inspector().evaluation()
```

```
     Evaluation(num_examples=59064655, accuracy=0.7076251440053277, loss=10.538257440461626, rmse=None, ndcg=None, aucs=None, auuc=None, qini=None)
```