RandomForest_GoldenEye

June 29, 2021

[21]:
```
##############################################################################
# Random Forest Classification Model (TensorFlow)                            #
# Based on the Implementation of:                                            #
# For GoldenEye Dataset                                                      #
# https://www.tensorflow.org/decision_forests/tutorials/beginner_colab       #
##############################################################################
```

[22]:
```
# Installieren aller benötigten Pakete
!pip install numpy==1.19.2
!pip install six==1.15.0
!pip install wheel==0.35
!pip install tensorflow_decision_forests
!pip install pandas
!pip install wurlitzer
!pip install matplotlib
!pip install onnxruntime
!pip install keras2onnx
```

```
Requirement already satisfied: numpy==1.19.2 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (1.19.2)
Requirement already satisfied: six==1.15.0 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (1.15.0)
Requirement already satisfied: wheel==0.35 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (0.35.0)
Requirement already satisfied: packaging>=20.2 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
wheel==0.35) (20.9)
Requirement already satisfied: pyparsing>=2.0.2 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
packaging>=20.2->wheel==0.35) (2.4.7)
Requirement already satisfied: tensorflow_decision_forests in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (0.1.7)
```

```
Requirement already satisfied: tensorflow~=2.5 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
tensorflow_decision_forests) (2.5.0)
Requirement already satisfied: absl-py in /home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
tensorflow_decision_forests) (0.13.0)
Requirement already satisfied: pandas in /home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
tensorflow_decision_forests) (1.2.5)
Requirement already satisfied: wheel in /home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
tensorflow_decision_forests) (0.35.0)
Requirement already satisfied: six in /home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
tensorflow_decision_forests) (1.15.0)
Requirement already satisfied: numpy in /home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
tensorflow_decision_forests) (1.19.2)
Requirement already satisfied: wrapt~=1.12.1 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
tensorflow~=2.5->tensorflow_decision_forests) (1.12.1)
Requirement already satisfied: grpcio~=1.34.0 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
tensorflow~=2.5->tensorflow_decision_forests) (1.34.1)
Requirement already satisfied: google-pasta~=0.2 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
tensorflow~=2.5->tensorflow_decision_forests) (0.2.0)
Requirement already satisfied: protobuf>=3.9.2 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
tensorflow~=2.5->tensorflow_decision_forests) (3.17.3)
Requirement already satisfied: tensorflow-estimator<2.6.0,>=2.5.0rc0 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
tensorflow~=2.5->tensorflow_decision_forests) (2.5.0)
Requirement already satisfied: termcolor~=1.1.0 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
tensorflow~=2.5->tensorflow_decision_forests) (1.1.0)
Requirement already satisfied: typing-extensions~=3.7.4 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
tensorflow~=2.5->tensorflow_decision_forests) (3.7.4.3)
Requirement already satisfied: keras-preprocessing~=1.1.2 in
```

```
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
tensorflow~=2.5->tensorflow_decision_forests) (1.1.2)
Requirement already satisfied: flatbuffers~=1.12.0 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
tensorflow~=2.5->tensorflow_decision_forests) (1.12)
Requirement already satisfied: gast==0.4.0 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
tensorflow~=2.5->tensorflow_decision_forests) (0.4.0)
Requirement already satisfied: tensorboard~=2.5 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
tensorflow~=2.5->tensorflow_decision_forests) (2.5.0)
Requirement already satisfied: h5py~=3.1.0 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
tensorflow~=2.5->tensorflow_decision_forests) (3.1.0)
Requirement already satisfied: keras-nightly~=2.5.0.dev in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
tensorflow~=2.5->tensorflow_decision_forests) (2.5.0.dev2021032900)
Requirement already satisfied: astunparse~=1.6.3 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
tensorflow~=2.5->tensorflow_decision_forests) (1.6.3)
Requirement already satisfied: opt-einsum~=3.3.0 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
tensorflow~=2.5->tensorflow_decision_forests) (3.3.0)
Requirement already satisfied: pytz>=2017.3 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
pandas->tensorflow_decision_forests) (2021.1)
Requirement already satisfied: python-dateutil>=2.7.3 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
pandas->tensorflow_decision_forests) (2.8.1)
Requirement already satisfied: packaging>=20.2 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
wheel->tensorflow_decision_forests) (20.9)
Requirement already satisfied: google-auth-oauthlib<0.5,>=0.4.1 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
tensorboard~=2.5->tensorflow~=2.5->tensorflow_decision_forests) (0.4.4)
Requirement already satisfied: google-auth<2,>=1.6.3 in
```

```
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
tensorboard~=2.5->tensorflow~=2.5->tensorflow_decision_forests) (1.32.0)
Requirement already satisfied: setuptools>=41.0.0 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
tensorboard~=2.5->tensorflow~=2.5->tensorflow_decision_forests) (44.0.0)
Requirement already satisfied: tensorboard-plugin-wit>=1.6.0 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
tensorboard~=2.5->tensorflow~=2.5->tensorflow_decision_forests) (1.8.0)
Requirement already satisfied: tensorboard-data-server<0.7.0,>=0.6.0 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
tensorboard~=2.5->tensorflow~=2.5->tensorflow_decision_forests) (0.6.1)
Requirement already satisfied: markdown>=2.6.8 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
tensorboard~=2.5->tensorflow~=2.5->tensorflow_decision_forests) (3.3.4)
Requirement already satisfied: requests<3,>=2.21.0 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
tensorboard~=2.5->tensorflow~=2.5->tensorflow_decision_forests) (2.25.1)
Requirement already satisfied: werkzeug>=0.11.15 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
tensorboard~=2.5->tensorflow~=2.5->tensorflow_decision_forests) (2.0.1)
Requirement already satisfied: pyparsing>=2.0.2 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
packaging>=20.2->wheel->tensorflow_decision_forests) (2.4.7)
Requirement already satisfied: requests-oauthlib>=0.7.0 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from google-
auth-oauthlib<0.5,>=0.4.1->tensorboard~=2.5->tensorflow~=2.5->tensorflow_decisio
n_forests) (1.3.0)
Requirement already satisfied: rsa<5,>=3.1.4; python_version >= "3.6" in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from google-
auth<2,>=1.6.3->tensorboard~=2.5->tensorflow~=2.5->tensorflow_decision_forests)
(4.7.2)
Requirement already satisfied: pyasn1-modules>=0.2.1 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from google-
auth<2,>=1.6.3->tensorboard~=2.5->tensorflow~=2.5->tensorflow_decision_forests)
(0.2.8)
Requirement already satisfied: cachetools<5.0,>=2.0.0 in
/home/julianbuecher/Projects/Bachelor-
```

```
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from google-
auth<2,>=1.6.3->tensorboard~=2.5->tensorflow~=2.5->tensorflow_decision_forests)
(4.2.2)
Requirement already satisfied: certifi>=2017.4.17 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from requests<3
,>=2.21.0->tensorboard~=2.5->tensorflow~=2.5->tensorflow_decision_forests)
(2021.5.30)
Requirement already satisfied: idna<3,>=2.5 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from requests<3
,>=2.21.0->tensorboard~=2.5->tensorflow~=2.5->tensorflow_decision_forests)
(2.10)
Requirement already satisfied: chardet<5,>=3.0.2 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from requests<3
,>=2.21.0->tensorboard~=2.5->tensorflow~=2.5->tensorflow_decision_forests)
(4.0.0)
Requirement already satisfied: urllib3<1.27,>=1.21.1 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from requests<3
,>=2.21.0->tensorboard~=2.5->tensorflow~=2.5->tensorflow_decision_forests)
(1.26.6)
Requirement already satisfied: oauthlib>=3.0.0 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from requests-
oauthlib>=0.7.0->google-auth-oauthlib<0.5,>=0.4.1->tensorboard~=2.5->tensorflow~
=2.5->tensorflow_decision_forests) (3.1.1)
Requirement already satisfied: pyasn1>=0.1.3 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
rsa<5,>=3.1.4; python_version >= "3.6"->google-
auth<2,>=1.6.3->tensorboard~=2.5->tensorflow~=2.5->tensorflow_decision_forests)
(0.4.8)
Requirement already satisfied: pandas in /home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (1.2.5)
Requirement already satisfied: numpy>=1.16.5 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from pandas)
(1.19.2)
Requirement already satisfied: pytz>=2017.3 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from pandas)
(2021.1)
Requirement already satisfied: python-dateutil>=2.7.3 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from pandas)
(2.8.1)
```

```
Requirement already satisfied: six>=1.5 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from python-
dateutil>=2.7.3->pandas) (1.15.0)
Requirement already satisfied: wurlitzer in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (2.1.0)
Requirement already satisfied: matplotlib in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (3.4.2)
Requirement already satisfied: numpy>=1.16 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
matplotlib) (1.19.2)
Requirement already satisfied: kiwisolver>=1.0.1 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
matplotlib) (1.3.1)
Requirement already satisfied: pyparsing>=2.2.1 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
matplotlib) (2.4.7)
Requirement already satisfied: python-dateutil>=2.7 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
matplotlib) (2.8.1)
Requirement already satisfied: pillow>=6.2.0 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
matplotlib) (8.2.0)
Requirement already satisfied: cycler>=0.10 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
matplotlib) (0.10.0)
Requirement already satisfied: six>=1.5 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from python-
dateutil>=2.7->matplotlib) (1.15.0)
Requirement already satisfied: onnxruntime in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (1.8.0)
Requirement already satisfied: numpy>=1.16.6 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
onnxruntime) (1.19.2)
Requirement already satisfied: flatbuffers in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
```

```
onnxruntime) (1.12)
Requirement already satisfied: protobuf in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
onnxruntime) (3.17.3)
Requirement already satisfied: six>=1.9 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
protobuf->onnxruntime) (1.15.0)
Requirement already satisfied: keras2onnx in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (1.7.0)
Requirement already satisfied: fire in /home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
keras2onnx) (0.4.0)
Requirement already satisfied: onnxconverter-common>=1.7.0 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
keras2onnx) (1.8.1)
Requirement already satisfied: requests in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
keras2onnx) (2.25.1)
Requirement already satisfied: onnx in /home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
keras2onnx) (1.9.0)
Requirement already satisfied: numpy in /home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
keras2onnx) (1.19.2)
Requirement already satisfied: protobuf in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
keras2onnx) (3.17.3)
Requirement already satisfied: termcolor in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
fire->keras2onnx) (1.1.0)
Requirement already satisfied: six in /home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
fire->keras2onnx) (1.15.0)
Requirement already satisfied: urllib3<1.27,>=1.21.1 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
requests->keras2onnx) (1.26.6)
Requirement already satisfied: idna<3,>=2.5 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
requests->keras2onnx) (2.10)
```

```
Requirement already satisfied: certifi>=2017.4.17 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
requests->keras2onnx) (2021.5.30)
Requirement already satisfied: chardet<5,>=3.0.2 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
requests->keras2onnx) (4.0.0)
Requirement already satisfied: typing-extensions>=3.6.2.1 in
/home/julianbuecher/Projects/Bachelor-
Thesis/ML.Proxy.Python.ModelTrainer/lib/python3.8/site-packages (from
onnx->keras2onnx) (3.7.4.3)
```

[23]:
```python
# Laden der benötigten Python Pakete
import os
# os.environ["TF_KERAS"]='1'
import pandas as pd
import numpy as np
import tensorflow_decision_forests as tfdf
import tensorflow as tf
from wurlitzer import sys_pipes
import matplotlib.pyplot as plt
import onnx
import keras2onnx as k2o
```

[24]:
```python
# Prüfung der installierten TensorFlow Decision Forests Version
print(f"Found TensorFlow Decision Forests v{tfdf.__version__}")
```

```
Found TensorFlow Decision Forests v0.1.7
```

[25]:
```python
# Laden der Netzwerk Traffic Daten für den GoldenEye Angriff
data_GoldenEye = pd.read_csv('../Data/Thursday-15-02-2018_GoldenEye-Attack.csv')
# Umbenennen der einzelnen Spalte für eine bessere Kompatibilität mit TensorFlow
data_GoldenEye.rename(columns={
    'Bwd Pkt Len Std':'bwd_pkt_len_std',
    'Flow IAT Min':'flow_iat_min',
    'Fwd IAT Min':'fwd_iat_min',
    'Flow IAT Mean':'flow_iat_mean',
    'Label':'label'
},
inplace=True)
```

[26]:
```python
# Festlegen des Wertes der bestimmten Variable
label = 'label'
```

[27]:
```python
# Aufteilen des Datasets in Training- und Test-Daten
def split_dataset(dataset,  test_ratio=0.30):
```

```
    """Splits a panda dataframe in two dataframes."""
    test_indices = np.random.rand(len(dataset)) < test_ratio
    return dataset[~test_indices], dataset[test_indices]

training_data_GoldenEye, testing_data_GoldenEye = split_dataset(data_GoldenEye)

print("{} examples in training, {} examples for testing.".format(
    len(training_data_GoldenEye), len(testing_data_GoldenEye)))
```

```
726398 examples in training, 311187 examples for testing.
```

[28]:
```
# Konvertieren des Panda Dataframes in ein TensorFlow Dataset
print("Converting Panda Dataframe into TensorFlow Dataset...")
training_dataset_GoldenEye = tfdf.keras.
 ↪pd_dataframe_to_tf_dataset(training_data_GoldenEye, label=label)
testing_dataset_GoldenEye = tfdf.keras.
 ↪pd_dataframe_to_tf_dataset(testing_data_GoldenEye, label=label)
```

```
Converting Panda Dataframe into TensorFlow Dataset…
```

[29]:
```
# Erstellen des Random Forest Modells
model = tfdf.keras.RandomForestModel()
model.compile(metrics=["accuracy"])
```

[30]:
```
# Trainieren des Modells
print("Training the Model: ")
with sys_pipes():
    model.fit(x=training_dataset_GoldenEye)
```

```
Training the Model:
11350/11350 [==============================] - 8s 693us/step
[INFO kernel.cc:746] Start Yggdrasil model training
[INFO kernel.cc:747] Collect training examples
[INFO kernel.cc:392] Number of batches: 11350
[INFO kernel.cc:393] Number of examples: 726398
[INFO kernel.cc:769] Dataset:
Number of records: 726398
Number of columns: 5

Number of columns by type:
        NUMERICAL: 4 (80%)
        CATEGORICAL: 1 (20%)

Columns:

NUMERICAL: 4 (80%)
        0: "bwd_pkt_len_std" NUMERICAL mean:121.803 min:0 max:898.733 sd:207.54
        1: "flow_iat_mean" NUMERICAL mean:2.95632e+06 min:0 max:1.19992e+08
```

```
sd:1.12199e+07
        2: "flow_iat_min" NUMERICAL mean:2.3833e+06 min:0 max:1.19992e+08
sd:1.10615e+07
        3: "fwd_iat_min" NUMERICAL mean:2.56421e+06 min:0 max:1.19992e+08
sd:1.15807e+07

CATEGORICAL: 1 (20%)
        4: "__LABEL" CATEGORICAL integerized vocab-size:3 no-ood-item

Terminology:
        nas: Number of non-available (i.e. missing) values.
        ood: Out of dictionary.
        manually-defined: Attribute which type is manually defined by the user
i.e. the type was not automatically inferred.
        tokenized: The attribute value is obtained through tokenization.
        has-dict: The attribute is attached to a string dictionary e.g. a
categorical attribute stored as a string.
        vocab-size: Number of unique values.

[INFO kernel.cc:772] Configure learner
[INFO kernel.cc:797] Training config:
learner: "RANDOM_FOREST"
features: "bwd_pkt_len_std"
features: "flow_iat_mean"
features: "flow_iat_min"
features: "fwd_iat_min"
label: "__LABEL"
task: CLASSIFICATION
[yggdrasil_decision_forests.model.random_forest.proto.random_forest_config] {
  num_trees: 300
  decision_tree {
    max_depth: 16
    min_examples: 5
    in_split_min_examples_check: true
    missing_value_policy: GLOBAL_IMPUTATION
    allow_na_conditions: false
    categorical_set_greedy_forward {
      sampling: 0.1
      max_num_items: -1
      min_item_frequency: 1
    }
    growing_strategy_local {
    }
    categorical {
      cart {
      }
    }
    num_candidate_attributes_ratio: -1
```

```
    axis_aligned_split {
    }
    internal {
      sorting_strategy: PRESORTED
    }
  }
  winner_take_all_inference: true
  compute_oob_performances: true
  compute_oob_variable_importances: false
  adapt_bootstrap_size_ratio_for_maximum_training_duration: false
}

[INFO kernel.cc:800] Deployment config:

[INFO kernel.cc:837] Train model
[INFO random_forest.cc:303] Training random forest on 726398 example(s) and 4
feature(s).
[INFO random_forest.cc:578] Training of tree  1/300 (tree index:5) done
accuracy:0.998265 logloss:0.0625336
[INFO random_forest.cc:578] Training of tree  11/300 (tree index:10) done
accuracy:0.998634 logloss:0.027996
[INFO random_forest.cc:578] Training of tree  21/300 (tree index:21) done
accuracy:0.998724 logloss:0.0215284
[INFO random_forest.cc:578] Training of tree  31/300 (tree index:30) done
accuracy:0.998727 logloss:0.0185809
[INFO random_forest.cc:578] Training of tree  41/300 (tree index:40) done
accuracy:0.998736 logloss:0.017375
[INFO random_forest.cc:578] Training of tree  51/300 (tree index:51) done
accuracy:0.998744 logloss:0.0162424
[INFO random_forest.cc:578] Training of tree  61/300 (tree index:60) done
accuracy:0.998754 logloss:0.0156105
[INFO random_forest.cc:578] Training of tree  71/300 (tree index:71) done
accuracy:0.998746 logloss:0.0150814
[INFO random_forest.cc:578] Training of tree  81/300 (tree index:80) done
accuracy:0.998756 logloss:0.0147768
[INFO random_forest.cc:578] Training of tree  91/300 (tree index:91) done
accuracy:0.99874 logloss:0.0145691
[INFO random_forest.cc:578] Training of tree  101/300 (tree index:100) done
accuracy:0.998751 logloss:0.0142987
[INFO random_forest.cc:578] Training of tree  111/300 (tree index:110) done
accuracy:0.99874 logloss:0.0140776
[INFO random_forest.cc:578] Training of tree  121/300 (tree index:120) done
accuracy:0.998732 logloss:0.0139975
[INFO random_forest.cc:578] Training of tree  131/300 (tree index:131) done
accuracy:0.998736 logloss:0.0139103
[INFO random_forest.cc:578] Training of tree  141/300 (tree index:141) done
accuracy:0.998744 logloss:0.0138692
[INFO random_forest.cc:578] Training of tree  151/300 (tree index:151) done
```

```
accuracy:0.998756 logloss:0.0135962
[INFO random_forest.cc:578] Training of tree  161/300 (tree index:160) done
accuracy:0.998753 logloss:0.0134597
[INFO random_forest.cc:578] Training of tree  171/300 (tree index:171) done
accuracy:0.998764 logloss:0.0132373
[INFO random_forest.cc:578] Training of tree  181/300 (tree index:180) done
accuracy:0.998772 logloss:0.0131511
[INFO random_forest.cc:578] Training of tree  191/300 (tree index:190) done
accuracy:0.998764 logloss:0.0129785
[INFO random_forest.cc:578] Training of tree  201/300 (tree index:199) done
accuracy:0.998768 logloss:0.0128869
[INFO random_forest.cc:578] Training of tree  211/300 (tree index:210) done
accuracy:0.998762 logloss:0.0128441
[INFO random_forest.cc:578] Training of tree  221/300 (tree index:220) done
accuracy:0.998765 logloss:0.0126692
[INFO random_forest.cc:578] Training of tree  231/300 (tree index:230) done
accuracy:0.998769 logloss:0.0126735
[INFO random_forest.cc:578] Training of tree  241/300 (tree index:240) done
accuracy:0.998765 logloss:0.0126322
[INFO random_forest.cc:578] Training of tree  251/300 (tree index:250) done
accuracy:0.998762 logloss:0.0125891
[INFO random_forest.cc:578] Training of tree  261/300 (tree index:260) done
accuracy:0.998758 logloss:0.01259
[INFO random_forest.cc:578] Training of tree  271/300 (tree index:270) done
accuracy:0.998761 logloss:0.0125446
[INFO random_forest.cc:578] Training of tree  281/300 (tree index:280) done
accuracy:0.998758 logloss:0.0124605
[INFO random_forest.cc:578] Training of tree  291/300 (tree index:290) done
accuracy:0.998765 logloss:0.0123756
[INFO random_forest.cc:578] Training of tree  300/300 (tree index:299) done
accuracy:0.998762 logloss:0.0123687
[INFO random_forest.cc:645] Final OOB metrics: accuracy:0.998762
logloss:0.0123687
[INFO kernel.cc:856] Export model in log directory: /tmp/tmpxbk0rqy7
[INFO kernel.cc:864] Save model in resources
[INFO kernel.cc:960] Loading model from path
[INFO decision_forest.cc:590] Model loaded with 300 root(s), 285772 node(s), and
4 input feature(s).
[INFO abstract_model.cc:973] Engine "RandomForestOptPred" built
[INFO kernel.cc:820] Use fast generic engine
```

[31]:
```python
# Evaluation des trainierten Modells mit den Testdaten
print("Evaluating the Model...")
evaluation = model.evaluate(testing_dataset_GoldenEye, return_dict=True)

print()
```

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

```
Evaluating the Model…
4863/4863 [==============================] - 16s 3ms/step - loss: 0.0000e+00 -
accuracy: 0.9988

loss: 0.0000
accuracy: 0.9988
```

[32]:
```
data_path = "../Data"
model_path = "Models"
onnx_path = "ONNX_Models"
model_name = "goldeneye_model"

# Trainiertes Modell für die spätere Verwendung abspeichern
model.save(os.path.join(data_path,model_path,model_name),overwrite=True)

# Konvertieren in das ONNX Modell
# onnx_model = k2o.convert_keras(model,df_model_name)
# onnx.save_model(onnx_model,os.path.join(data_path,onnx_path,model_name + ".
 ↪onnx"))
```

```
INFO:tensorflow:Assets written to: ../Data/Models/goldeneye_model/assets
INFO:tensorflow:Assets written to: ../Data/Models/goldeneye_model/assets
```

[33]:
```
# Plotten des ersten Baumes innerhalb des Decision Forests
with open('../Data/Models/GoldenEye_Model_Tree.html', 'w') as f:
    f.write(tfdf.model_plotter.plot_model(model, tree_idx=0, max_depth=3))
tfdf.model_plotter.plot_model(model, tree_idx=0, max_depth=3)
```

[33]:
```
'\n<script src="https://d3js.org/d3.v6.min.js"></script>\n<div
id="tree_plot_f07efa5daa9a4200b90dffb44454e813"></div>\n<script>\n/*\n *
Copyright 2021 Google LLC.\n * Licensed under the Apache License, Version 2.0
(the "License");\n * you may not use this file except in compliance with the
License.\n * You may obtain a copy of the License at\n *\n *
https://www.apache.org/licenses/LICENSE-2.0\n *\n * Unless required by
applicable law or agreed to in writing, software\n * distributed under the
License is distributed on an "AS IS" BASIS,\n * WITHOUT WARRANTIES OR CONDITIONS
OF ANY KIND, either express or implied.\n * See the License for the specific
language governing permissions and\n * limitations under the License.\n
*/\n\n/**\n *  Plotting of decision trees generated by TF-DF.\n *\n *  A tree is
a recursive structure of node objects.\n *  A node contains one or more of the
following components:\n *\n *    - A value: Representing the output of the node.
If the node is not a leaf,\n *      the value is only present for analysis i.e.
it is not used for\n *      predictions.\n *\n *    - A condition : For non-leaf
nodes, the condition (also known as split)\n *      defines a binary test to
```

```
branch to the positive or negative child.\n *\n *    - An explanation: Generally
a plot showing the relation between the label\n *      and the condition to give
insights about the effect of the condition.\n *\n *    - Two children : For non-
leaf nodes, the children nodes. The first\n *      children (i.e.
"node.children[0]") is the negative children (drawn in\n *      red). The second
children is the positive one (drawn in green).\n *\n */\n\n/**\n * Plots a
single decision tree into a DOM element.\n * @param {!options} options
Dictionary of configurations.\n * @param {!tree} raw_tree Recursive tree
structure.\n * @param {string} canvas_id Id of the output dom element.\n
*/\nfunction display_tree(options, raw_tree, canvas_id) {\n
console.log(options);\n\n  // Determine the node placement.\n  const tree_struct
= d3.tree().nodeSize(\n      [options.node_y_offset,
options.node_x_offset])(d3.hierarchy(raw_tree));\n\n  // Boundaries of the node
placement.\n  let x_min = Infinity;\n  let x_max = -x_min;\n  let y_min =
Infinity;\n  let y_max = -x_min;\n\n  tree_struct.each(d => {\n    if (d.x >
x_max) x_max = d.x;\n    if (d.x < x_min) x_min = d.x;\n    if (d.y > y_max)
y_max = d.y;\n    if (d.y < y_min) y_min = d.y;\n  });\n\n  // Size of the
plot.\n  const width = y_max - y_min + options.node_x_size + options.margin *
2;\n  const height = x_max - x_min + options.node_y_size + options.margin * 2
+\n      options.node_y_offset - options.node_y_size;\n\n  const plot =
d3.select(canvas_id);\n\n  // Tool tip\n  options.tooltip =
plot.append(\'div\')\n                          .attr(\'width\', 100)\n
.attr(\'height\', 100)\n                            .style(\'padding\', \'4px\')\n
.style(\'background\', \'#fff\')\n                              .style(\'box-shadow\',
\'4px 4px 0px rgba(0,0,0,0.1)\')\n                              .style(\'border\',
\'1px solid black\')\n                          .style(\'font-family\', \'sans-
serif\')\n                       .style(\'font-size\', options.font_size)\n
.style(\'position\', \'absolute\')\n                           .style(\'z-index\',
\'10\')\n                        .attr(\'pointer-events\', \'none\')\n
.style(\'display\', \'none\');\n\n  // Create canvas\n  const svg =
plot.append(\'svg\').attr(\'width\', width).attr(\'height\', height);\n  const
graph =\n      svg.style(\'overflow\', \'visible\')\n          .append(\'g\')\n
.attr(\'font-family\', \'sans-serif\')\n          .attr(\'font-size\',
options.font_size)\n          .attr(\n              \'transform\',\n
() => `translate(${options.margin},${\n                  - x_min +
options.node_y_offset / 2 + options.margin})`);\n\n  // Plot bounding box.\n  if
(options.show_plot_bounding_box) {\n    svg.append(\'rect\')\n
.attr(\'width\', width)\n        .attr(\'height\', height)\n
.attr(\'fill\', \'none\')\n        .attr(\'stroke-width\', 1.0)\n
.attr(\'stroke\', \'black\');\n  }\n\n  // Draw the edges.\n
display_edges(options, graph, tree_struct);\n\n  // Draw the nodes.\n
display_nodes(options, graph, tree_struct);\n}\n\n/**\n * Draw the nodes of the
tree.\n * @param {!options} options Dictionary of configurations.\n * @param
{!graph} graph D3 search handle containing the graph.\n * @param {!tree_struct}
tree_struct Structure of the tree (node placement,\n *     data, etc.).\n
*/\nfunction display_nodes(options, graph, tree_struct) {\n  const nodes =
graph.append(\'g\')\n                    .selectAll(\'g\')\n
```

```
.data(tree_struct.descendants())\n                    .join(\'g\')\n
.attr(\'transform\', d => `translate(${d.y},${d.x})`);\n\n
nodes.append(\'rect\')\n      .attr(\'x\', 0.5)\n      .attr(\'y\', 0.5)\n
.attr(\'width\', options.node_x_size)\n      .attr(\'height\',
options.node_y_size)\n      .attr(\'stroke\', \'lightgrey\')\n
.attr(\'stroke-width\', 1)\n      .attr(\'fill\', \'white\')\n      .attr(\'y\',
-options.node_y_size / 2);\n\n  // Brackets on the right of condition nodes
without children.\n  non_leaf_node_without_children =\n      nodes.filter(node
=> node.data.condition != null && node.children == null)\n
.append(\'g\')\n          .attr(\'transform\',
`translate(${options.node_x_size},0)`);\n\n
non_leaf_node_without_children.append(\'path\')\n      .attr(\'d\', \'M0,0 C
10,0 0,10 10,10\')\n      .attr(\'fill\', \'none\')\n      .attr(\'stroke-
width\', 1.0)\n      .attr(\'stroke\', \'#F00\');\n\n
non_leaf_node_without_children.append(\'path\')\n      .attr(\'d\', \'M0,0 C
10,0 0,-10 10,-10\')\n      .attr(\'fill\', \'none\')\n      .attr(\'stroke-
width\', 1.0)\n      .attr(\'stroke\', \'#0F0\');\n\n  const node_content =
nodes.append(\'g\').attr(\n      \'transform\',\n
`translate(0,${options.node_padding - options.node_y_size / 2})`);\n\n
node_content.append(node => create_node_element(options, node));\n}\n\n/**\n *
Creates the D3 content for a single node.\n * @param {!options} options
Dictionary of configurations.\n * @param {!node} node Node to draw.\n * @return
{!d3} D3 content.\n */\nfunction create_node_element(options, node) {\n  //
Output accumulator.\n  let output = {\n    // Content to draw.\n    content:
d3.create(\'svg:g\'),\n    // Vertical offset to the next element to draw.\n
vertical_offset: 0\n  };\n\n  // Conditions.\n  if (node.data.condition != null)
{\n    display_condition(options, node.data.condition, output);\n  }\n\n  //
Values.\n  if (node.data.value != null) {\n    display_value(options,
node.data.value, output);\n  }\n\n  // Explanations.\n  if
(node.data.explanation != null) {\n    display_explanation(options,
node.data.explanation, output);\n  }\n\n  return
output.content.node();\n}\n\n\n/**\n * Adds a single line of text inside of a
node.\n * @param {!options} options Dictionary of configurations.\n * @param
{string} text Text to display.\n * @param {!output} output Output display
accumulator.\n */\nfunction display_node_text(options, text, output) {\n
output.content.append(\'text\')\n      .attr(\'x\', options.node_padding)\n
.attr(\'y\', output.vertical_offset)\n      .attr(\'alignment-baseline\',
\'hanging\')\n      .text(text);\n  output.vertical_offset += 10;\n}\n\n/**\n *
Adds a single line of text inside of a node with a tooltip.\n * @param
{!options} options Dictionary of configurations.\n * @param {string} text Text
to display.\n * @param {string} tooltip Text in the Tooltip.\n * @param
{!output} output Output display accumulator.\n */\nfunction
display_node_text_with_tooltip(options, text, tooltip, output) {\n  const item =
output.content.append(\'text\')\n                    .attr(\'x\',
options.node_padding)\n                    .attr(\'alignment-baseline\',
\'hanging\')\n                    .text(text);\n\n  add_tooltip(options, item, ()
=> tooltip);\n  output.vertical_offset += 10;\n}\n\n/**\n * Adds a tooltip to a
```

```
dom element.\n * @param {!options} options Dictionary of configurations.\n *
@param {!dom} target Dom element to equip with a tooltip.\n * @param {!func}
get_content Generates the html content of the tooltip.\n */\nfunction
add_tooltip(options, target, get_content) {\n  function show(d) {\n
options.tooltip.style(\'display\', \'block\');\n
options.tooltip.html(get_content());\n  }\n\n  function hide(d) {\n
options.tooltip.style(\'display\', \'none\');\n  }\n\n  function move(d) {\n
options.tooltip.style(\'display\', \'block\');\n
options.tooltip.style(\'left\', (d.pageX + 5) + \'px\');\n
options.tooltip.style(\'top\', d.pageY + \'px\');\n  }\n\n
target.on(\'mouseover\', show);\n  target.on(\'mouseout\', hide);\n
target.on(\'mousemove\', move);\n}\n\n/**\n * Adds a condition inside of a
node.\n * @param {!options} options Dictionary of configurations.\n * @param
{!condition} condition Condition to display.\n * @param {!output} output Output
display accumulator.\n */\nfunction display_condition(options, condition,
output) {\n  threshold_format = d3.format(\'r\');\n\n  if (condition.type ===
\'IS_MISSING\') {\n    display_node_text(options, `${condition.attribute} is
missing`, output);\n    return;\n  }\n\n  if (condition.type === \'IS_TRUE\')
{\n    display_node_text(options, `${condition.attribute} is true`, output);\n
return;\n  }\n\n  if (condition.type === \'NUMERICAL_IS_HIGHER_THAN\') {\n
format = d3.format(\'r\');\n    display_node_text(\n        options,\n
`${condition.attribute} >= ${threshold_format(condition.threshold)}`,\n
output);\n    return;\n  }\n\n  if (condition.type === \'CATEGORICAL_IS_IN\')
{\n    display_node_text_with_tooltip(\n        options, `${condition.attribute}
in […]`,\n        `${condition.attribute} in [${condition.mask}]`, output);\n
return;\n  }\n\n  if (condition.type === \'CATEGORICAL_SET_CONTAINS\') {\n
display_node_text_with_tooltip(\n        options, `${condition.attribute}
intersect […]`,\n        `${condition.attribute} intersect
[${condition.mask}]`, output);\n    return;\n  }\n\n  if (condition.type ===
\'NUMERICAL_SPARSE_OBLIQUE\') {\n    display_node_text_with_tooltip(\n
options, `Sparse oblique split…`,\n
`[${condition.attributes}]*[${condition.weights}]>=${\n
threshold_format(condition.threshold)}`,\n        output);\n    return;\n  }\n\n
display_node_text(\n      options, `Non supported condition ${condition.type}`,
output);\n}\n\n/**\n * Adds a value inside of a node.\n * @param {!options}
options Dictionary of configurations.\n * @param {!value} value Value to
display.\n * @param {!output} output Output display accumulator.\n */\nfunction
display_value(options, value, output) {\n  if (value.type === \'PROBABILITY\')
{\n    const left_margin = 0;\n    const right_margin = 50;\n    const
plot_width = options.node_x_size - options.node_padding * 2 -\n
left_margin - right_margin;\n\n    let cusum =
Array.from(d3.cumsum(value.distribution));\n    cusum.unshift(0);\n    const
distribution_plot = output.content.append(\'g\').attr(\n        \'transform\',
`translate(0,${output.vertical_offset + 0.5})`);\n\n
distribution_plot.selectAll(\'rect\')\n        .data(value.distribution)\n
.join(\'rect\')\n        .attr(\'height\', 10)\n        .attr(\n
\'x\',\n            (d, i) =>\n                (cusum[i] * plot_width +
```

```
left_margin + options.node_padding))\n        .attr(\'width\', (d, i) => d *
plot_width)\n        .style(\'fill\', (d, i) => d3.schemeSet1[i]);\n\n    const
num_examples =\n        output.content.append(\'g\')\n
.attr(\'transform\', `translate(0,${output.vertical_offset})`)\n
.append(\'text\')\n            .attr(\'x\', options.node_x_size -
options.node_padding)\n            .attr(\'alignment-baseline\', \'hanging\')\n
.attr(\'text-anchor\', \'end\')\n
.text(`(${value.num_examples})`);\n\n    const distribution_details =
d3.create(\'ul\');\n    distribution_details.selectAll(\'li\')\n
.data(value.distribution)\n        .join(\'li\')\n        .append(\'span\')\n
.text(\n            (d, i) =>\n                \'class \' + i + \': \' +
d3.format(\'.3%\')(value.distribution[i]));\n\n    add_tooltip(options,
distribution_plot, () => distribution_details.html());\n    add_tooltip(options,
num_examples, () => \'Number of examples\');\n\n    output.vertical_offset +=
10;\n    return;\n  }\n\n  if (value.type === \'REGRESSION\') {\n
display_node_text(\n        options,\n        \'value: \' +
d3.format(\'r\')(value.value) + ` (` +\n
d3.format(\'.6\')(value.num_examples) + `)`,\n        output);\n    return;\n
}\n\n  display_node_text(options, `Non supported value ${value.type}`,
output);\n}\n\n/**\n * Adds an explanation inside of a node.\n * @param
{!options} options Dictionary of configurations.\n * @param {!explanation}
explanation Explanation to display.\n * @param {!output} output Output display
accumulator.\n */\nfunction display_explanation(options, explanation, output)
{\n  // Margin before the explanation.\n  output.vertical_offset += 10;\n\n
display_node_text(\n      options, `Non supported explanation
${explanation.type}`, output);\n}\n\n\n/**\n * Draw the edges of the tree.\n *
@param {!options} options Dictionary of configurations.\n * @param {!graph}
graph D3 search handle containing the graph.\n * @param {!tree_struct}
tree_struct Structure of the tree (node placement,\n *     data, etc.).\n
*/\nfunction display_edges(options, graph, tree_struct) {\n  // Draw an edge
between a parent and a child node with a bezier.\n  function draw_single_edge(d)
{\n    return \'M\' + (d.source.y + options.node_x_size) + \',\' + d.source.x +
\' C\' +\n        (d.source.y + options.node_x_size + options.edge_rounding) +
\',\' +\n        d.source.x + \' \' + (d.target.y - options.edge_rounding) +
\',\' +\n        d.target.x + \' \' + d.target.y + \',\' + d.target.x;\n  }\n\n
graph.append(\'g\')\n      .attr(\'fill\', \'none\')\n      .attr(\'stroke-
width\', 1.2)\n      .selectAll(\'path\')\n      .data(tree_struct.links())\n
.join(\'path\')\n      .attr(\'d\', draw_single_edge)\n      .attr(\n
\'stroke\', d => (d.target === d.source.children[0]) ? \'#0F0\' :
\'#F00\');\n}\n\ndisplay_tree({"margin": 10, "node_x_size": 160, "node_y_size":
28, "node_x_offset": 180, "node_y_offset": 33, "font_size": 10, "edge_rounding":
20, "node_padding": 2, "show_plot_bounding_box": false}, {"value": {"type":
"PROBABILITY", "distribution": [0.9602339213489024, 0.03976607865109761],
"num_examples": 726398.0}, "condition": {"type": "NUMERICAL_IS_HIGHER_THAN",
"attribute": "flow_iat_mean", "threshold": 605939.75}, "children": [{"value":
{"type": "PROBABILITY", "distribution": [0.847308997956201, 0.152691002043799],
"num_examples": 171739.0}, "condition": {"type": "NUMERICAL_IS_HIGHER_THAN",
```

"attribute": "flow_iat_mean", "threshold": 2140764.0}, "children": [{"value":
{"type": "PROBABILITY", "distribution": [0.9418933129884323,
0.05810668701156772], "num_examples": 96994.0}, "condition": {"type":
"NUMERICAL_IS_HIGHER_THAN", "attribute": "fwd_iat_min", "threshold": 2141266.5},
"children": [{"value": {"type": "PROBABILITY", "distribution":
[0.86626072922016, 0.13373927077984007], "num_examples": 40893.0}, "condition":
{"type": "NUMERICAL_IS_HIGHER_THAN", "attribute": "flow_iat_min", "threshold":
52046848.0}}, {"value": {"type": "PROBABILITY", "distribution":
[0.9970232259674515, 0.0029767740325484394], "num_examples": 56101.0},
"condition": {"type": "NUMERICAL_IS_HIGHER_THAN", "attribute":
"bwd_pkt_len_std", "threshold": 486.000732421875}}]}, {"value": {"type":
"PROBABILITY", "distribution": [0.7245702053649073, 0.27542979463509265],
"num_examples": 74745.0}, "condition": {"type": "NUMERICAL_IS_HIGHER_THAN",
"attribute": "bwd_pkt_len_std", "threshold": 330.9661865234375}, "children":
[{"value": {"type": "PROBABILITY", "distribution": [0.5001726573536337,
0.49982734264636636], "num_examples": 31855.0}, "condition": {"type":
"NUMERICAL_IS_HIGHER_THAN", "attribute": "flow_iat_min", "threshold": 1.5}},
{"value": {"type": "PROBABILITY", "distribution": [0.891233387736069,
0.10876661226393099], "num_examples": 42890.0}, "condition": {"type":
"NUMERICAL_IS_HIGHER_THAN", "attribute": "flow_iat_min", "threshold":
1357616.0}}]}]}, {"value": {"type": "PROBABILITY", "distribution":
[0.9951988519072078, 0.004801148092792148], "num_examples": 554659.0},
"condition": {"type": "NUMERICAL_IS_HIGHER_THAN", "attribute": "flow_iat_min",
"threshold": 22.5}, "children": [{"value": {"type": "PROBABILITY",
"distribution": [0.9999848466098078, 1.5153390192220755e-05], "num_examples":
395951.0}, "condition": {"type": "NUMERICAL_IS_HIGHER_THAN", "attribute":
"fwd_iat_min", "threshold": 49688.5}, "children": [{"value": {"type":
"PROBABILITY", "distribution": [0.9997831586555836, 0.00021684134441633538],
"num_examples": 27670.0}, "condition": {"type": "NUMERICAL_IS_HIGHER_THAN",
"attribute": "flow_iat_min", "threshold": 49680.0}}, {"value": {"type":
"PROBABILITY", "distribution": [1.0, 0.0], "num_examples": 368281.0}}]},
{"value": {"type": "PROBABILITY", "distribution": [0.9832585628953802,
0.016741437104619804], "num_examples": 158708.0}, "condition": {"type":
"NUMERICAL_IS_HIGHER_THAN", "attribute": "fwd_iat_min", "threshold": 24.5},
"children": [{"value": {"type": "PROBABILITY", "distribution":
[0.9556101494664592, 0.04438985053354087], "num_examples": 53323.0},
"condition": {"type": "NUMERICAL_IS_HIGHER_THAN", "attribute": "flow_iat_min",
"threshold": 1.5}}, {"value": {"type": "PROBABILITY", "distribution":
[0.9972481852256013, 0.0027518147743986338], "num_examples": 105385.0},
"condition": {"type": "NUMERICAL_IS_HIGHER_THAN", "attribute":
"bwd_pkt_len_std", "threshold": 381.69085693359375}}]}]}]},
"#tree_plot_f07efa5daa9a4200b90dffb44454e813")\n</script>\n'

[34]:
```
# Erstellen einer Bilanz für das trainierte Modell
model.summary()
```

```
Model: "random_forest_model_1"
```

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
Total params: 1
Trainable params: 0
Non-trainable params: 1
_________________________________________________________________
Type: "RANDOM_FOREST"
Task: CLASSIFICATION
Label: "__LABEL"

Input Features (4):
        bwd_pkt_len_std
        flow_iat_mean
        flow_iat_min
        fwd_iat_min

No weights

Variable Importance: NUM_NODES:
    1.   "flow_iat_mean" 46884.000000 ################
    2.     "fwd_iat_min" 38819.000000 ###########
    3.    "flow_iat_min" 37415.000000 ##########
    4. "bwd_pkt_len_std" 19618.000000

Variable Importance: NUM_AS_ROOT:
    1.   "flow_iat_mean" 214.000000 ################
    2. "bwd_pkt_len_std" 58.000000 ##
    3.     "fwd_iat_min" 28.000000

Variable Importance: SUM_SCORE:
    1.   "flow_iat_mean" 13013827.343974 ################
    2.    "flow_iat_min" 10860271.115623 ###########
    3. "bwd_pkt_len_std" 7855573.952822 ######
    4.     "fwd_iat_min" 4455481.723066

Variable Importance: MEAN_MIN_DEPTH:
    1.         "__LABEL" 11.915364 ################
    2. "bwd_pkt_len_std"  4.938714 #####
    3.     "fwd_iat_min"  4.019469 ####
    4.    "flow_iat_min"  2.515558 ##
    5.   "flow_iat_mean"  0.803955



Winner take all: true
Out-of-bag evaluation: accuracy:0.998762 logloss:0.0123687
Number of trees: 300
```

```
Total number of nodes: 285772

Number of nodes by tree:
Count: 300 Average: 952.573 StdDev: 82.246
Min: 729 Max: 1121 Ignored: 0
---------------------------------------------
[  729,  748)  5   1.67%   1.67% ##
[  748,  768)  5   1.67%   3.33% ##
[  768,  787)  3   1.00%   4.33% #
[  787,  807)  4   1.33%   5.67% #
[  807,  827)  3   1.00%   6.67% #
[  827,  846)  7   2.33%   9.00% ##
[  846,  866) 18   6.00%  15.00% ######
[  866,  886) 19   6.33%  21.33% ######
[  886,  905) 16   5.33%  26.67% #####
[  905,  925) 22   7.33%  34.00% #######
[  925,  945) 31  10.33%  44.33% ##########
[  945,  964) 28   9.33%  53.67% #########
[  964,  984) 24   8.00%  61.67% ########
[  984, 1004) 29   9.67%  71.33% #########
[ 1004, 1023) 23   7.67%  79.00% #######
[ 1023, 1043) 18   6.00%  85.00% ######
[ 1043, 1063) 18   6.00%  91.00% ######
[ 1063, 1082) 12   4.00%  95.00% ####
[ 1082, 1102) 11   3.67%  98.67% ####
[ 1102, 1121]  4   1.33% 100.00% #

Depth by leafs:
Count: 143036 Average: 11.9091 StdDev: 2.65945
Min: 2 Max: 15 Ignored: 0
---------------------------------------------
[  2,  3)    39   0.03%   0.03%
[  3,  4)   148   0.10%   0.13%
[  4,  5)   644   0.45%   0.58%
[  5,  6)  2276   1.59%   2.17% #
[  6,  7)  3452   2.41%   4.59% #
[  7,  8)  4461   3.12%   7.70% ##
[  8,  9)  6321   4.42%  12.12% ##
[  9, 10)  8853   6.19%  18.31% ###
[ 10, 11) 12751   8.91%  27.23% ####
[ 11, 12) 16334  11.42%  38.65% ######
[ 12, 13) 18982  13.27%  51.92% #######
[ 13, 14) 20212  14.13%  66.05% #######
[ 14, 15) 19405  13.57%  79.61% #######
[ 15, 15] 29158  20.39% 100.00% ##########

Number of training obs by leaf:
Count: 143036 Average: 1523.53 StdDev: 17670
```

```
Min: 5 Max: 494588 Ignored: 0
---------------------------------------------
[      5,  24734) 141456  98.90%  98.90% ##########
[  24734,  49463)    980   0.69%  99.58%
[  49463,  74192)    147   0.10%  99.68%
[  74192,  98921)     55   0.04%  99.72%
[  98921, 123651)     17   0.01%  99.73%
[ 123651, 148380)     42   0.03%  99.76%
[ 148380, 173109)     11   0.01%  99.77%
[ 173109, 197838)     11   0.01%  99.78%
[ 197838, 222567)     19   0.01%  99.79%
[ 222567, 247297)     11   0.01%  99.80%
[ 247297, 272026)     26   0.02%  99.82%
[ 272026, 296755)     37   0.03%  99.84%
[ 296755, 321484)     22   0.02%  99.86%
[ 321484, 346213)     33   0.02%  99.88%
[ 346213, 370943)     20   0.01%  99.90%
[ 370943, 395672)     44   0.03%  99.93%
[ 395672, 420401)     37   0.03%  99.95%
[ 420401, 445130)     34   0.02%  99.98%
[ 445130, 469859)     25   0.02%  99.99%
[ 469859, 494588]      9   0.01% 100.00%

Attribute in nodes:
        46884 : flow_iat_mean [NUMERICAL]
        38819 : fwd_iat_min [NUMERICAL]
        37415 : flow_iat_min [NUMERICAL]
        19618 : bwd_pkt_len_std [NUMERICAL]

Attribute in nodes with depth <= 0:
        214 : flow_iat_mean [NUMERICAL]
        58 : bwd_pkt_len_std [NUMERICAL]
        28 : fwd_iat_min [NUMERICAL]

Attribute in nodes with depth <= 1:
        377 : flow_iat_mean [NUMERICAL]
        271 : bwd_pkt_len_std [NUMERICAL]
        178 : flow_iat_min [NUMERICAL]
        74 : fwd_iat_min [NUMERICAL]

Attribute in nodes with depth <= 2:
        680 : bwd_pkt_len_std [NUMERICAL]
        621 : flow_iat_mean [NUMERICAL]
        491 : flow_iat_min [NUMERICAL]
        269 : fwd_iat_min [NUMERICAL]

Attribute in nodes with depth <= 3:
        1381 : flow_iat_min [NUMERICAL]
```

```
        1151 : flow_iat_mean [NUMERICAL]
        1118 : bwd_pkt_len_std [NUMERICAL]
        585 : fwd_iat_min [NUMERICAL]

Attribute in nodes with depth <= 5:
        3750 : flow_iat_mean [NUMERICAL]
        3681 : flow_iat_min [NUMERICAL]
        3190 : bwd_pkt_len_std [NUMERICAL]
        2450 : fwd_iat_min [NUMERICAL]

Condition type in nodes:
        142736 : HigherCondition
Condition type in nodes with depth <= 0:
        300 : HigherCondition
Condition type in nodes with depth <= 1:
        900 : HigherCondition
Condition type in nodes with depth <= 2:
        2061 : HigherCondition
Condition type in nodes with depth <= 3:
        4235 : HigherCondition
Condition type in nodes with depth <= 5:
        13071 : HigherCondition
Node format: NOT_SET

Training OOB:
        trees: 1, Out-of-bag evaluation: accuracy:0.998265 logloss:0.0625336
        trees: 11, Out-of-bag evaluation: accuracy:0.998634 logloss:0.027996
        trees: 21, Out-of-bag evaluation: accuracy:0.998724 logloss:0.0215284
        trees: 31, Out-of-bag evaluation: accuracy:0.998727 logloss:0.0185809
        trees: 41, Out-of-bag evaluation: accuracy:0.998736 logloss:0.017375
        trees: 51, Out-of-bag evaluation: accuracy:0.998744 logloss:0.0162424
        trees: 61, Out-of-bag evaluation: accuracy:0.998754 logloss:0.0156105
        trees: 71, Out-of-bag evaluation: accuracy:0.998746 logloss:0.0150814
        trees: 81, Out-of-bag evaluation: accuracy:0.998756 logloss:0.0147768
        trees: 91, Out-of-bag evaluation: accuracy:0.99874 logloss:0.0145691
        trees: 101, Out-of-bag evaluation: accuracy:0.998751 logloss:0.0142987
        trees: 111, Out-of-bag evaluation: accuracy:0.99874 logloss:0.0140776
        trees: 121, Out-of-bag evaluation: accuracy:0.998732 logloss:0.0139975
        trees: 131, Out-of-bag evaluation: accuracy:0.998736 logloss:0.0139103
        trees: 141, Out-of-bag evaluation: accuracy:0.998744 logloss:0.0138692
        trees: 151, Out-of-bag evaluation: accuracy:0.998756 logloss:0.0135962
        trees: 161, Out-of-bag evaluation: accuracy:0.998753 logloss:0.0134597
        trees: 171, Out-of-bag evaluation: accuracy:0.998764 logloss:0.0132373
        trees: 181, Out-of-bag evaluation: accuracy:0.998772 logloss:0.0131511
        trees: 191, Out-of-bag evaluation: accuracy:0.998764 logloss:0.0129785
        trees: 201, Out-of-bag evaluation: accuracy:0.998768 logloss:0.0128869
        trees: 211, Out-of-bag evaluation: accuracy:0.998762 logloss:0.0128441
        trees: 221, Out-of-bag evaluation: accuracy:0.998765 logloss:0.0126692
```

```
trees: 231, Out-of-bag evaluation: accuracy:0.998769 logloss:0.0126735
trees: 241, Out-of-bag evaluation: accuracy:0.998765 logloss:0.0126322
trees: 251, Out-of-bag evaluation: accuracy:0.998762 logloss:0.0125891
trees: 261, Out-of-bag evaluation: accuracy:0.998758 logloss:0.01259
trees: 271, Out-of-bag evaluation: accuracy:0.998761 logloss:0.0125446
trees: 281, Out-of-bag evaluation: accuracy:0.998758 logloss:0.0124605
trees: 291, Out-of-bag evaluation: accuracy:0.998765 logloss:0.0123756
trees: 300, Out-of-bag evaluation: accuracy:0.998762 logloss:0.0123687
```

[35]:
```
# Erstellen von Grafiken für die Effizienz des Trainings
logs = model.make_inspector().training_logs()
plt.figure(figsize=(12,4))

plt.subplot(1,2,1)
plt.plot([log.num_trees for log in logs], [log.evaluation.accuracy for log in␣
↪logs])
plt.xlabel("Number of trees")
plt.ylabel("Accuracy (out-of-bag)")

plt.subplot(1,2,2)
plt.plot([log.num_trees for log in logs], [log.evaluation.loss for log in logs])
plt.xlabel("Number of trees")
plt.ylabel("Logloss (out-of-bag)")

plt.savefig('../Data/Visualized/GoldenEye_Model.png')
plt.clf()
```

```
<Figure size 864x288 with 0 Axes>
```