In [ ]:
```
pip install h2o
```

```
Looking in indexes: https://pypi.org/simple, (https://pypi.org/simple,) https://us-python.pkg.dev/colab-wheels/public/simple/ (https://us-python.pkg.dev/colab-wheels/public/simple/)
Collecting h2o
  Downloading h2o-3.36.1.2.tar.gz (177.0 MB)
     |████████████████████████████████| 177.0 MB 31 kB/s
Requirement already satisfied: requests in /usr/local/lib/python3.7/dist-packages (from h2o) (2.23.0)
Requirement already satisfied: tabulate in /usr/local/lib/python3.7/dist-packages (from h2o) (0.8.9)
Requirement already satisfied: future in /usr/local/lib/python3.7/dist-packages (from h2o) (0.16.0)
Requirement already satisfied: chardet<4,>=3.0.2 in /usr/local/lib/python3.7/dist-packages (from requests->h2o) (3.0.4)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.7/dist-packages (from requests->h2o) (2022.5.18.1)
Requirement already satisfied: urllib3!=1.25.0,!=1.25.1,<1.26,>=1.21.1 in /usr/local/lib/python3.7/dist-packages (from requests->h2o) (1.24.3)
Requirement already satisfied: idna<3,>=2.5 in /usr/local/lib/python3.7/dist-packages (from requests->h2o) (2.10)
Building wheels for collected packages: h2o
  Building wheel for h2o (setup.py) ... done
  Created wheel for h2o: filename=h2o-3.36.1.2-py2.py3-none-any.whl size=177070945 sha256=440e05334859f1b887534340551b73c847fc3f437b7759e642013f6200ecdaec
  Stored in directory: /root/.cache/pip/wheels/a5/2b/ae/534595df09001e1573b1046be62a3f049507307340119db65c
Successfully built h2o
Installing collected packages: h2o
Successfully installed h2o-3.36.1.2
```

In [ ]:
```
!nvidia-smi
```

```
NVIDIA-SMI has failed because it couldn't communicate with the NVIDIA driver. Make sure that the latest NVIDIA driver is installed and running.
```

In [ ]:
```
!apt-get install default-jre
!java -version
```

```
Reading package lists... Done
Building dependency tree
Reading state information... Done
default-jre is already the newest version (2:1.11-68ubuntu1~18.04.1).
default-jre set to manually installed.
The following package was automatically installed and is no longer required:
  libnvidia-common-460
Use 'apt autoremove' to remove it.
0 upgraded, 0 newly installed, 0 to remove and 45 not upgraded.
openjdk version "11.0.15" 2022-04-19
OpenJDK Runtime Environment (build 11.0.15+10-Ubuntu-0ubuntu0.18.04.1)
OpenJDK 64-Bit Server VM (build 11.0.15+10-Ubuntu-0ubuntu0.18.04.1, mixed mode, sharing)
```

In [ ]:
```
!pip install h2o
```

```
Looking in indexes: https://pypi.org/simple, (https://pypi.org/simple,) https://us-python.pkg.dev/colab-wheels/public/simple/ (https://us-python.pkg.dev/colab-wheels/public/simple/)
Requirement already satisfied: h2o in /usr/local/lib/python3.7/dist-packages (3.36.1.2)
Requirement already satisfied: future in /usr/local/lib/python3.7/dist-packages (from h2o) (0.16.0)
Requirement already satisfied: tabulate in /usr/local/lib/python3.7/dist-packages (from h2o) (0.8.9)
Requirement already satisfied: requests in /usr/local/lib/python3.7/dist-packages (from h2o) (2.23.0)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.7/dist-packages (from requests->h2o) (2022.5.18.1)
Requirement already satisfied: urllib3!=1.25.0,!=1.25.1,<1.26,>=1.21.1 in /usr/local/lib/python3.7/dist-packages (from requests->h2o) (1.24.3)
Requirement already satisfied: chardet<4,>=3.0.2 in /usr/local/lib/python3.7/dist-packages (from requests->h2o) (3.0.4)
Requirement already satisfied: idna<3,>=2.5 in /usr/local/lib/python3.7/dist-packages (from requests->h2o) (2.10)
```

In [ ]:
```
import h2o
```

In [ ]:

```
h2o.init()
```

```
Checking whether there is an H2O instance running at http://localhost:54321 (http://localhost:54321) ..... not f
ound.
Attempting to start a local H2O server...
  Java Version: openjdk version "11.0.15" 2022-04-19; OpenJDK Runtime Environment (build 11.0.15+10-Ubuntu-0ubun
tu0.18.04.1); OpenJDK 64-Bit Server VM (build 11.0.15+10-Ubuntu-0ubuntu0.18.04.1, mixed mode, sharing)
  Starting server from /usr/local/lib/python3.7/dist-packages/h2o/backend/bin/h2o.jar
  Ice root: /tmp/tmponsss4kk
  JVM stdout: /tmp/tmponsss4kk/h2o_unknownUser_started_from_python.out
  JVM stderr: /tmp/tmponsss4kk/h2o_unknownUser_started_from_python.err
  Server is running at http://127.0.0.1:54321 (http://127.0.0.1:54321)
Connecting to H2O server at http://127.0.0.1:54321 (http://127.0.0.1:54321) ... successful.
```

H2O_cluster_uptime:	03 secs
H2O_cluster_timezone:	Etc/UTC
H2O_data_parsing_timezone:	UTC
H2O_cluster_version:	3.36.1.2
H2O_cluster_version_age:	7 days, 15 hours and 38 minutes
H2O_cluster_name:	H2O_from_python_unknownUser_d2ux6d
H2O_cluster_total_nodes:	1
H2O_cluster_free_memory:	3.172 Gb
H2O_cluster_total_cores:	2
H2O_cluster_allowed_cores:	2
H2O_cluster_status:	locked, healthy
H2O_connection_url:	http://127.0.0.1:54321
H2O_connection_proxy:	{"http": null, "https": null}
H2O_internal_security:	False
Python_version:	3.7.13 final

In [ ]:
```
from h2o.automl import H2OAutoML
```

In [ ]:
```
df = h2o.import_file('/content/drive/MyDrive/IIIT-K Intern/country_wise_latest.csv')
```

```
Parse progress: |████████████████████████████████████████████████████████████| (done) 100%
```

In [ ]:
```
df.types
```

Out[18]:
```
{'1 week % increase': 'real',
 '1 week change': 'int',
 'Active': 'int',
 'Confirmed': 'int',
 'Confirmed last week': 'int',
 'Country/Region': 'string',
 'Deaths': 'int',
 'Deaths / 100 Cases': 'real',
 'Deaths / 100 Recovered': 'real',
 'New cases': 'int',
 'New deaths': 'int',
 'New recovered': 'int',
 'Recovered': 'int',
 'Recovered / 100 Cases': 'real',
 'WHO Region': 'enum'}
```

In [ ]:
```
df.describe()
```

```
Rows:187
Cols:15
```

	Country/Region	Confirmed	Deaths	Recovered	Active	New cases	New
type	string	int	int	int	int	int	
mins	NaN	10.0	0.0	0.0	0.0	0.0	
mean	NaN	88130.93582887705	3497.518716577539	50631.48128342246	34001.93582887701	1222.9572192513367	28.957219251
maxs	NaN	4290259.0	148011.0	1846641.0	2816444.0	56336.0	
sigma	NaN	383318.6638306154	14100.00248201848	190188.18964313966	213326.17337142891	5710.374790280563	120.037172955
zeros	0	0	17	6	5	33	
missing	0	0	0	0	0	0	
0	Afghanistan	36263.0	1269.0	25198.0	9796.0	106.0	
1	Albania	4880.0	144.0	2745.0	1991.0	117.0	
2	Algeria	27973.0	1163.0	18837.0	7973.0	616.0	
3	Andorra	907.0	52.0	803.0	52.0	10.0	
4	Angola	950.0	41.0	242.0	667.0	18.0	
5	Antigua and Barbuda	86.0	3.0	65.0	18.0	4.0	
6	Argentina	167416.0	3059.0	72575.0	91782.0	4890.0	
7	Armenia	37390.0	711.0	26665.0	10014.0	73.0	
8	Australia	15303.0	167.0	9311.0	5825.0	368.0	
9	Austria	20558.0	713.0	18246.0	1599.0	86.0	

In [ ]:
```
df_train,df_test,df_valid = df.split_frame(ratios=[.7, .15])
```

In [ ]:
```
df_train
```

Out[21]:

Country/Region	Confirmed	Deaths	Recovered	Active	New cases	New deaths	New recovered	Deaths / 100 Cases	Recovered / 100 Cases	Deaths / 100 Recovered	Confirmed last week	1 week change	1 w incre
Algeria	27973	1163	18837	7973	616	8	749	4.16	67.34	6.17	23691	4282	18
Andorra	907	52	803	52	10	0	0	5.73	88.53	6.48	884	23	
Angola	950	41	242	667	18	1	0	4.32	25.47	16.94	749	201	26
Antigua and Barbuda	86	3	65	18	4	0	5	3.49	75.58	4.62	76	10	13
Austria	20558	713	18246	1599	86	1	37	3.47	88.75	3.91	19743	815	4
Azerbaijan	30446	423	23242	6781	396	6	558	1.39	76.34	1.82	27890	2556	9
Bahamas	382	11	91	280	40	0	0	2.88	23.82	12.09	174	208	119
Bangladesh	226225	2965	125683	97577	2772	37	1801	1.31	55.56	2.36	207453	18772	9
Barbados	110	7	94	9	0	0	0	6.36	85.45	7.45	106	4	3
Belarus	67251	538	60492	6221	119	4	67	0.8	89.95	0.89	66213	1038	1

In [ ]:
```
y = "WHO Region"
x = df.columns
x.remove(y)
x.remove("Country/Region")
```

In [ ]:
```
aml = H2OAutoML(max_models = 10, seed = 10, exclude_algos = ["StackedEnsemble", "DeepLearning"], verbosity="info"
```

In [ ]:
```
aml.train(x = x, y = y, training_frame = df_train, validation_frame=df_valid)
```

```
AutoML progress: |
06:45:38.605: Project: AutoML_2_20220603_64538
06:45:38.605: Cross-validation disabled by user: no fold column nor nfolds > 1.
06:45:38.609: Setting stopping tolerance adaptively based on the training frame: 0.05
06:45:38.609: Build control seed: 10
06:45:38.610: training frame: Frame key: AutoML_2_20220603_64538_training_py_2_sid_9a32    cols: 15    rows:
135  chunks: 1    size: 9556  checksum: -6096892916768307408
06:45:38.610: validation frame: Frame key: py_4_sid_9a32    cols: 15    rows: 28  chunks: 1    size: 3498  ch
ecksum: -9067743764230477440
06:45:38.610: leaderboard frame: Frame key: py_4_sid_9a32    cols: 15    rows: 28  chunks: 1    size: 3498  c
hecksum: -9067743764230477440
06:45:38.610: blending frame: NULL
06:45:38.610: response column: WHO Region
06:45:38.610: fold column: null
06:45:38.610: weights column: null
06:45:38.611: Loading execution steps: [{XGBoost : [def_2 (1g, 10w), def_1 (2g, 10w), def_3 (3g, 10w), grid_1
(4g, 90w), lr_search (7g, 30w)]}, {GLM : [def_1 (1g, 10w)]}, {DRF : [def_1 (2g, 10w), XRT (3g, 10w)]}, {GBM :
[def_5 (1g, 10w), def_2 (2g, 10w), def_3 (2g, 10w), def_4 (2g, 10w), def_1 (3g, 10w), grid_1 (4g, 60w), lr_an
```

In [ ]:
```
lb = aml.leaderboard
```

In [ ]:
```
lb.head()
```

model_id	mean_per_class_error	logloss	rmse	mse
GBM_5_AutoML_2_20220603_64538	0.614815	1.82064	0.709052	0.502755
GBM_3_AutoML_2_20220603_64538	0.685185	1.8582	0.73148	0.535063
DRF_1_AutoML_2_20220603_64538	0.685185	2.04599	0.750693	0.563541
XGBoost_2_AutoML_2_20220603_64538	0.688889	1.73555	0.725417	0.52623
XGBoost_1_AutoML_2_20220603_64538	0.731481	1.69912	0.763917	0.58357
GBM_4_AutoML_2_20220603_64538	0.740741	1.85461	0.737587	0.544035
GLM_1_AutoML_2_20220603_64538	0.740741	1.67038	0.775431	0.601294
XRT_1_AutoML_2_20220603_64538	0.759259	3.04748	0.745781	0.556189
GBM_2_AutoML_2_20220603_64538	0.759259	1.89026	0.746182	0.556787
XGBoost_3_AutoML_2_20220603_64538	0.814815	1.96622	0.75793	0.574458

Out[44]:

In [ ]:
```
df_pred=aml.leader.predict(df_test)
```

```
gbm prediction progress: |████████████████████████████████████████████████| (done) 100%
```

In [ ]:
```
df_pred.head()
```

Out[46]:

predict	Africa	Americas	Eastern Mediterranean	Europe	South-East Asia	Western Pacific
South-East Asia	0.0529445	0.0160772	0.0479261	0.0241309	0.828224	0.0306974
Europe	0.137384	0.0302537	0.0451859	0.742155	0.0230462	0.0219751
Europe	0.184158	0.0226405	0.243165	0.336711	0.199883	0.0134433
Europe	0.00428107	0.00364757	0.0109031	0.97349	0.00480615	0.00287265
Americas	0.100839	0.42125	0.0551942	0.091431	0.306877	0.0244099
Africa	0.610431	0.141922	0.0436069	0.158654	0.0232681	0.022117
South-East Asia	0.0986816	0.0163821	0.0275585	0.343741	0.498779	0.0148583
Eastern Mediterranean	0.124357	0.101823	0.376713	0.0957351	0.283764	0.0176093
Europe	0.0716283	0.357699	0.0827013	0.416687	0.0365349	0.0347495
Europe	0.0140248	0.193718	0.0122824	0.761374	0.00836107	0.010239

In [ ]:
```
aml.leader.model_performance(df_test)
```

```
ModelMetricsMultinomial: gbm
** Reported on test data. **

MSE: 0.5064778787990475
RMSE: 0.7116725924180638
LogLoss: 1.753464831031331
Mean Per-Class Error: 0.5105820105820106
AUC: NaN
AUCPR: NaN
Multinomial auc values: Table is not computed because it is disabled (model parameter 'auc_type' is set to AUTO
or NONE) or due to domain size (maximum is 50 domains).
Multinomial auc_pr values: Table is not computed because it is disabled (model parameter 'auc_type' is set to AU
TO or NONE) or due to domain size (maximum is 50 domains).

Confusion Matrix: Row labels: Actual class; Column labels: Predicted class
```

	Africa	Americas	Eastern Mediterranean	Europe	South-East Asia	Western Pacific	Error	Rate
0	1.0	1.0	0.0	0.0	0.0	0.0	0.500000	1 / 2
1	2.0	2.0	1.0	3.0	1.0	0.0	0.777778	7 / 9
2	0.0	1.0	0.0	2.0	0.0	1.0	1.000000	4 / 4
3	0.0	1.0	0.0	5.0	1.0	0.0	0.285714	2 / 7
4	0.0	0.0	0.0	0.0	0.0	0.0	NaN	0 / 0
5	0.0	0.0	0.0	1.0	0.0	1.0	0.500000	1 / 2
6	3.0	5.0	1.0	11.0	2.0	2.0	0.625000	15 / 24

```
Top-6 Hit Ratios:
```

	k	hit_ratio
0	1	0.375000
1	2	0.708333
2	3	0.750000
3	4	0.833333
4	5	0.916667
5	6	1.000000

Out[47]:

In [ ]:
```
model_ids = list(aml.leaderboard['model_id'].as_data_frame().iloc[:,0])
```

In [ ]:
```
model_ids
```

Out[49]:
```
['GBM_5_AutoML_2_20220603_64538',
 'GBM_3_AutoML_2_20220603_64538',
 'DRF_1_AutoML_2_20220603_64538',
 'XGBoost_2_AutoML_2_20220603_64538',
 'XGBoost_1_AutoML_2_20220603_64538',
 'GBM_4_AutoML_2_20220603_64538',
 'GLM_1_AutoML_2_20220603_64538',
 'XRT_1_AutoML_2_20220603_64538',
 'GBM_2_AutoML_2_20220603_64538',
 'XGBoost_3_AutoML_2_20220603_64538']
```

In [ ]:
```
h2o.get_model([mid for mid in model_ids if "XGBoost" in mid][0])
```

```
Model Details
=============
H2OXGBoostEstimator :  XGBoost
Model Key:  XGBoost_2_AutoML_2_20220603_64538


Model Summary:
```

	number_of_trees
0	30.0

```


ModelMetricsMultinomial: xgboost
** Reported on train data. **

MSE: 0.305689706208584
```

In [ ]:
```
out = h2o.get_model([mid for mid in model_ids if "XGBoost" in mid][0])
```

In [ ]:
```
out.params
```

Out[52]:
```
{'auc_type': {'actual': 'AUTO', 'default': 'AUTO', 'input': 'AUTO'},
 'backend': {'actual': 'cpu', 'default': 'auto', 'input': 'auto'},
 'booster': {'actual': 'gbtree', 'default': 'gbtree', 'input': 'gbtree'},
 'build_tree_one_node': {'actual': False, 'default': False, 'input': False},
 'calibrate_model': {'actual': False, 'default': False, 'input': False},
 'calibration_frame': {'actual': None, 'default': None, 'input': None},
 'categorical_encoding': {'actual': 'OneHotInternal',
  'default': 'AUTO',
  'input': 'AUTO'},
 'checkpoint': {'actual': None, 'default': None, 'input': None},
 'col_sample_rate': {'actual': 0.8, 'default': 1.0, 'input': 0.8},
 'col_sample_rate_per_tree': {'actual': 0.8, 'default': 1.0, 'input': 0.8},
 'colsample_bylevel': {'actual': 0.8, 'default': 1.0, 'input': 1.0},
 'colsample_bynode': {'actual': 1.0, 'default': 1.0, 'input': 1.0},
 'colsample_bytree': {'actual': 0.8, 'default': 1.0, 'input': 1.0},
 'distribution': {'actual': 'multinomial',
  'default': 'AUTO',
  'input': 'multinomial'},
 'dmatrix_type': {'actual': 'dense', 'default': 'auto', 'input': 'auto'},
```

In [ ]:
```
out.convert_H2OXGBoostParams_2_XGBoostParams()
```

Out[53]:
```
({'alpha': 0.0,
  'booster': 'gbtree',
  'colsample_bylevel': 0.8,
  'colsample_bytree': 0.8,
  'eta': 0.3,
  'gamma': 0.0,
  'grow_policy': 'depthwise',
  'lambda': 1.0,
  'max_delta_step': 0.0,
  'max_depth': 10,
  'min_child_weight': 5.0,
  'nround': 10000,
  'nthread': 2,
  'num_class': 6,
  'objective': 'multi:softprob',
  'seed': 13,
  'silent': True,
  'subsample': 0.6,
  'tree_method': 'exact'},
 10000)
```

In [ ]:
```
out
```

```
Model Details
=============
H2OXGBoostEstimator :  XGBoost
Model Key:  XGBoost_2_AutoML_2_20220603_64538


Model Summary:
```

	number_of_trees
0	30.0

```
ModelMetricsMultinomial: xgboost
** Reported on train data. **

MSE: 0.305689706208584
```

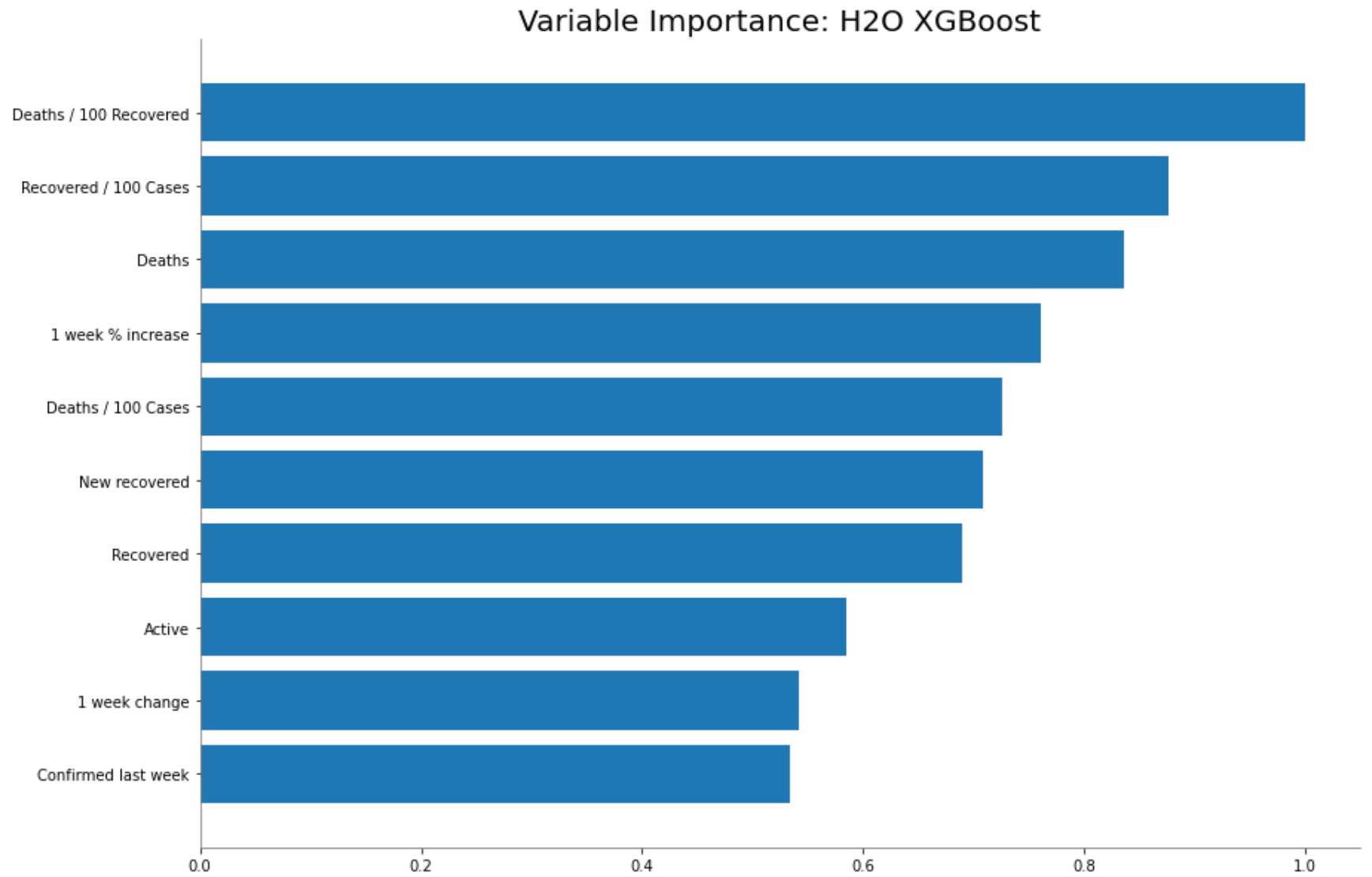
In [ ]:
```
out_gbm = h2o.get_model([mid for mid in model_ids if "GBM" in mid][0])
```

In [ ]:
```
out.confusion_matrix()
```

```
---------------------------------------------------------------------------
TypeError                                 Traceback (most recent call last)
<ipython-input-56-ad6d01e05f24> in <module>()
----> 1 out.confusion_matrix()

TypeError: confusion_matrix() missing 1 required positional argument: 'data'
```

In [ ]:
```
out.varimp_plot()
```



Out[57]:
```
<h2o.plot._plot_result._MObject at 0x7ff9170f34d0>

<Figure size 432x288 with 0 Axes>
```

In [ ]: