Hewlett Packard
Enterprise

SMOKE TEST

H2O 3.24

Date Prepared: Sept 2019

Hewlett Packard
Enterprise

Document Information

Project Name	**EPIC Accelerator Deployment & Integration Services**		
Project Owner		Document Version No	0.1
Quality Review Method			
Prepared By		Preparation Date	Sept 2019
Reviewed By		Review Date	

Table of Contents

1 CREATING A FRESH FLOW FILE FROM H2O 4
1.1 ACCESSING WEBPAGE OF H2O 4
1.2 CREATE A NEW FLOW FILE 4
1.3 CREATE A MODEL 8

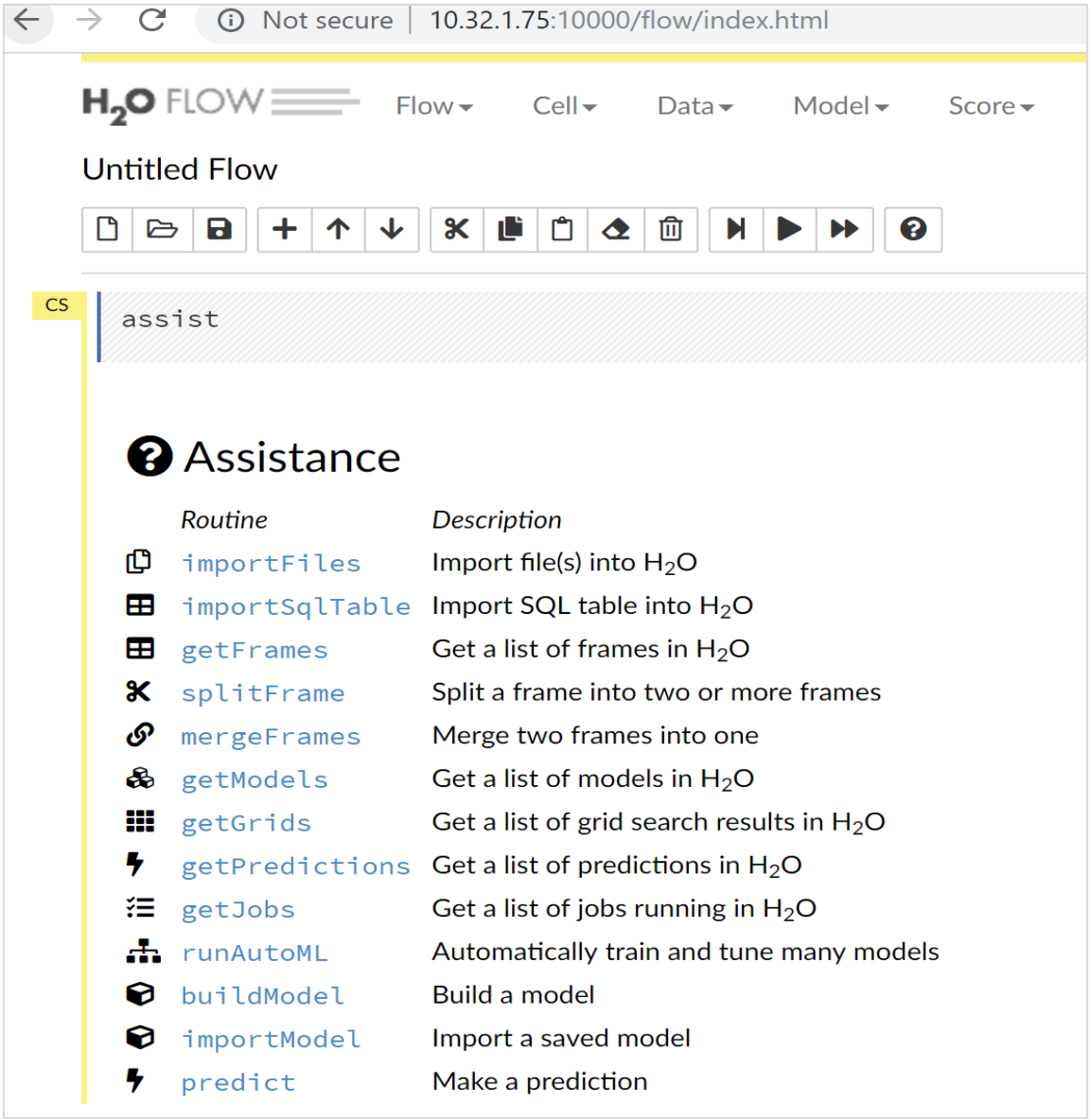
Table of Tables

NO TABLE OF FIGURES ENTRIES FOUND.

1 CREATING A FRESH FLOW FILE FROM H2O

1.1 Accessing Webpage of H2O

From EPIC cluster page, click on master. A new window will pop-up with H2O dashboard.



Flow is a home page for H2O.ai. Flow is designed to help data scientists to automate manuals tasks like file importing, get frames, models etc.

There is a help section in the right corner of panel, which can be used as a reference.

1.2 Create a new Flow File

Below is the source code, to create a new flow file:

```
# One has the option to download the files as well
# Point to folder containing them
location = "../../bigdata/laptop/kdd2009/small-churn/"

# For this example we will use public files available at S3 on AWS
hasLocalData = false

# Configure file locations
if hasLocalData
 trainFile = location + "kdd_train.csv"
 validFile = location + "kdd_valid.csv"
else
 trainFile = "https://h2o-public-test-
data.s3.amazonaws.com/bigdata/laptop/kdd2009/small-
churn/kdd_train.csv"
 validFile = "https://h2o-public-test-
data.s3.amazonaws.com/bigdata/laptop/kdd2009/small-
churn/kdd_valid.csv"

parseFiles

  paths: [trainFile]
  destination_frame: "kdd_train.hex"
  parse_type: "CSV"
  separator: 44
  number_columns: 231
  single_quotes: false
  column_names:

["churn","Var1","Var2","Var3","Var4","Var5","Var6","Var7","Var8","Va
r9","Var10","Var11","Var12","Var13","Var14","Var15","Var16","Var17",
"Var18","Var19","Var20","Var21","Var22","Var23","Var24","Var25","Var
26","Var27","Var28","Var29","Var30","Var31","Var32","Var33","Var34",
"Var35","Var36","Var37","Var38","Var39","Var40","Var41","Var42","Var
43","Var44","Var45","Var46","Var47","Var48","Var49","Var50","Var51",
"Var52","Var53","Var54","Var55","Var56","Var57","Var58","Var59","Var
60","Var61","Var62","Var63","Var64","Var65","Var66","Var67","Var68",
"Var69","Var70","Var71","Var72","Var73","Var74","Var75","Var76","Var
77","Var78","Var79","Var80","Var81","Var82","Var83","Var84","Var85",
"Var86","Var87","Var88","Var89","Var90","Var91","Var92","Var93","Var
94","Var95","Var96","Var97","Var98","Var99","Var100","Var101","Var10
2","Var103","Var104","Var105","Var106","Var107","Var108","Var109","V
ar110","Var111","Var112","Var113","Var114","Var115","Var116","Var117
","Var118","Var119","Var120","Var121","Var122","Var123","Var124","Va
r125","Var126","Var127","Var128","Var129","Var130","Var131","Var132"
,"Var133","Var134","Var135","Var136","Var137","Var138","Var139","Var
140","Var141","Var142","Var143","Var144","Var145","Var146","Var147",
```

```
"Var148","Var149","Var150","Var151","Var152","Var153","Var154","Var1
55","Var156","Var157","Var158","Var159","Var160","Var161","Var162","
Var163","Var164","Var165","Var166","Var167","Var168","Var169","Var17
0","Var171","Var172","Var173","Var174","Var175","Var176","Var177","V
ar178","Var179","Var180","Var181","Var182","Var183","Var184","Var185
","Var186","Var187","Var188","Var189","Var190","Var191","Var192","Va
r193","Var194","Var195","Var196","Var197","Var198","Var199","Var200"
,"Var201","Var202","Var203","Var204","Var205","Var206","Var207","Var
208","Var209","Var210","Var211","Var212","Var213","Var214","Var215",
"Var216","Var217","Var218","Var219","Var220","Var221","Var222","Var2
23","Var224","Var225","Var226","Var227","Var228","Var229","Var230"]
  column_types:

["Enum","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric",
"Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric
","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numer
ic","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Num
eric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","N
umeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric",
"Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric
","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numer
ic","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Num
eric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","N
umeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric",
"Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric
","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numer
ic","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Num
eric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","N
umeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric",
"Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric
","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numer
ic","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Num
eric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","N
umeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric",
"Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric
","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numer
ic","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Num
eric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","N
umeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric",
"Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric
","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numer
ic","Enum","Enum","Enum","Enum","Enum","Enum","Enum","Enum","Enum","
Enum","Enum","Enum","Enum","Enum","Enum","Enum","Enum","Enum","Numer
ic","Enum","Enum","Enum","Enum","Enum","Enum","Enum","Enum","Enum","
Enum","Enum","Enum","Enum","Enum","Enum","Enum","Enum","Enum","Enum"
,"Enum","Numeric"]
  delete_on_done: true
  check_header: 1
```

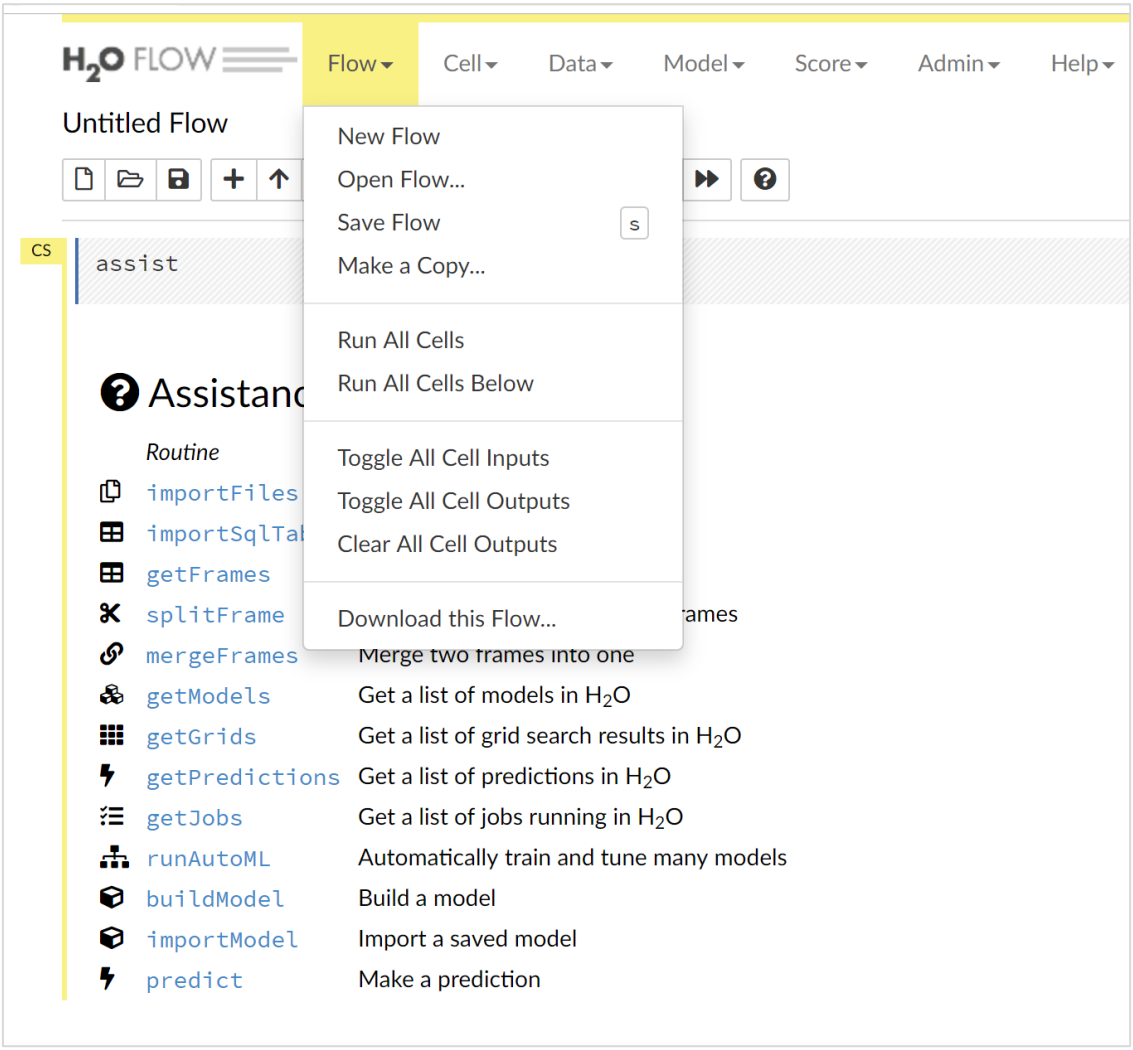
```
  chunk_size: 4194304

  parseFiles
  paths: [validFile]
  destination_frame: "kdd_valid.hex"
  parse_type: "CSV"
  separator: 44
  number_columns: 231
  single_quotes: false
  column_names:

["churn","Var1","Var2","Var3","Var4","Var5","Var6","Var7","Var8","Va
r9","Var10","Var11","Var12","Var13","Var14","Var15","Var16","Var17",
"Var18","Var19","Var20","Var21","Var22","Var23","Var24","Var25","Var
26","Var27","Var28","Var29","Var30","Var31","Var32","Var33","Var34",
"Var35","Var36","Var37","Var38","Var39","Var40","Var41","Var42","Var
43","Var44","Var45","Var46","Var47","Var48","Var49","Var50","Var51",
"Var52","Var53","Var54","Var55","Var56","Var57","Var58","Var59","Var
60","Var61","Var62","Var63","Var64","Var65","Var66","Var67","Var68",
"Var69","Var70","Var71","Var72","Var73","Var74","Var75","Var76","Var
77","Var78","Var79","Var80","Var81","Var82","Var83","Var84","Var85",
"Var86","Var87","Var88","Var89","Var90","Var91","Var92","Var93","Var
94","Var95","Var96","Var97","Var98","Var99","Var100","Var101","Var10
2","Var103","Var104","Var105","Var106","Var107","Var108","Var109","V
ar110","Var111","Var112","Var113","Var114","Var115","Var116","Var117
","Var118","Var119","Var120","Var121","Var122","Var123","Var124","Va
r125","Var126","Var127","Var128","Var129","Var130","Var131","Var132"
,"Var133","Var134","Var135","Var136","Var137","Var138","Var139","Var
140","Var141","Var142","Var143","Var144","Var145","Var146","Var147",
"Var148","Var149","Var150","Var151","Var152","Var153","Var154","Var1
55","Var156","Var157","Var158","Var159","Var160","Var161","Var162","
Var163","Var164","Var165","Var166","Var167","Var168","Var169","Var17
0","Var171","Var172","Var173","Var174","Var175","Var176","Var177","V
ar178","Var179","Var180","Var181","Var182","Var183","Var184","Var185
","Var186","Var187","Var188","Var189","Var190","Var191","Var192","Va
r193","Var194","Var195","Var196","Var197","Var198","Var199","Var200"
,"Var201","Var202","Var203","Var204","Var205","Var206","Var207","Var
208","Var209","Var210","Var211","Var212","Var213","Var214","Var215",
"Var216","Var217","Var218","Var219","Var220","Var221","Var222","Var2
23","Var224","Var225","Var226","Var227","Var228","Var229","Var230"]
  column_types:

["Enum","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric",
"Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric
","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numer
ic","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Num
eric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","N
umeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric",
```

```
"Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric
","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numer
ic","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Num
eric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","N
umeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric",
"Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric
","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numer
ic","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Num
eric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","N
umeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric",
"Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric
","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numer
ic","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Num
eric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","N
umeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric",
"Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric
","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numer
ic","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Num
eric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","N
umeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric",
"Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric
","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numer
ic","Enum","Enum","Enum","Enum","Enum","Enum","Enum","Enum","Enum","
Enum","Enum","Enum","Enum","Enum","Enum","Enum","Enum","Enum","Numer
ic","Enum","Enum","Enum","Enum","Enum","Enum","Enum","Enum","Enum","
Enum","Enum","Enum","Enum","Enum","Enum","Enum","Enum","Enum","Enum"
,"Enum","Numeric"]
  delete_on_done: true
  check_header: 1
  chunk_size: 4194304


buildModel 'gbm', {"model_id":"gbm-
model","training_frame":"kdd_train.hex","validation_frame":"kdd_vali
d.hex","ignore_const_cols":true,"score_each_iteration":false,"respon
se_column":"churn","ntrees":50,"max_depth":"3","min_rows":10,"nbins"
:"4","nbins_cats":"20","r2_stopping":0.999999,"learn_rate":"0.12","d
istribution":"bernoulli","balance_classes":false,"seed":-
1941834077389261000}

getModel "gbm-model"
```

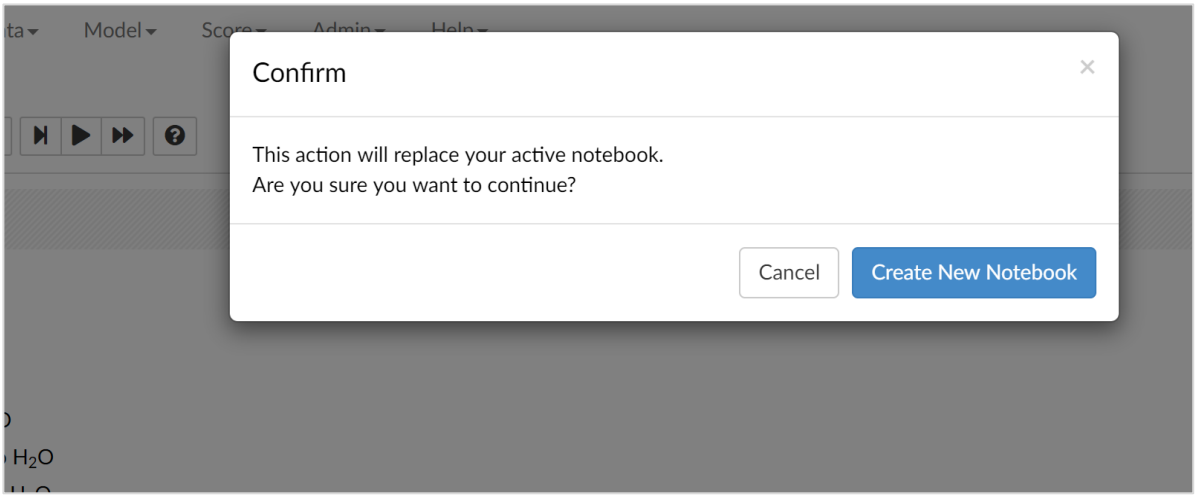
1.3 Create a Model

In this section, we will create a Model using Flow.

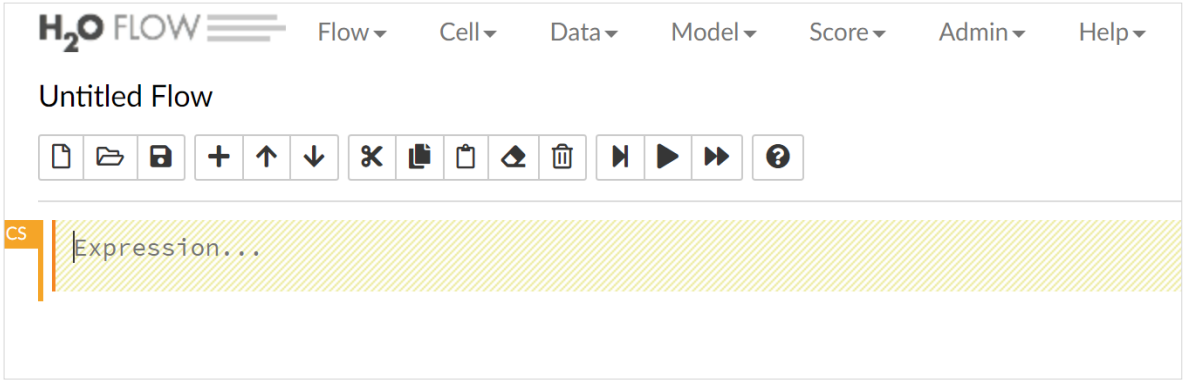
1. From the Menu bar, click on **Flow** drop-down menu and select **New Flow**



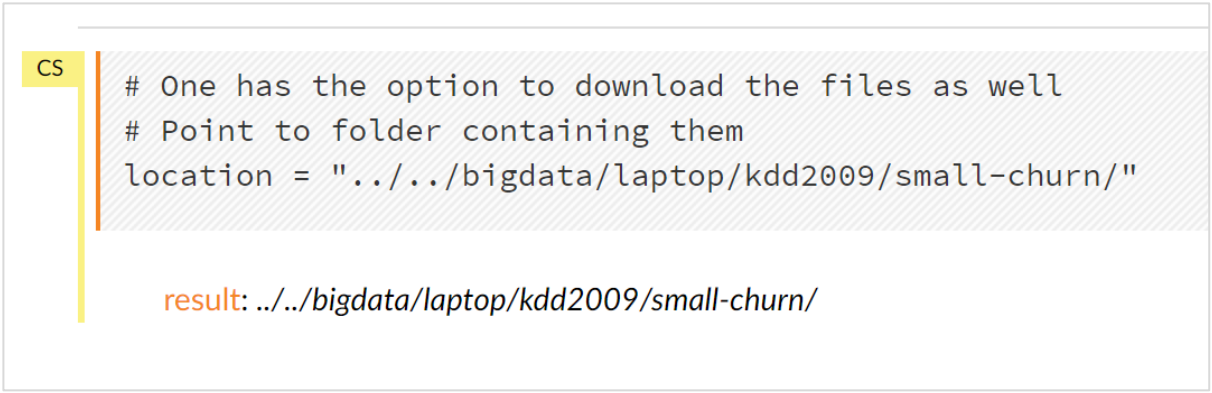
2. A Confirm dialog box will appear. Click on **Create New Notebook**



3. A new flow will look like below:



4. To execute any code, enter the code in the code cell and press **Alt+Enter**



5. To add a new cell, click on **+**

6. Add the remaining code blocks and evaluate the same (Use code provided in Section: 1.2)

```
result: ../../bigdata/laptop/kdd2009/small-churn/
```

```
# For this example we will use public files available at S3 on AWS
hasLocalData = false

# Configure file locations
if hasLocalData
 trainFile = location + "kdd_train.csv"
 validFile = location + "kdd_valid.csv"
else
 trainFile = "https://h2o-public-test-data.s3.amazonaws.com/bigdata/laptop/kdd2009/small-churn/kdd_train.csv"
 validFile = "https://h2o-public-test-data.s3.amazonaws.com/bigdata/laptop/kdd2009/small-churn/kdd_valid.csv"
```

```
result: https://h2o-public-test-data.s3.amazonaws.com/bigdata/laptop/kdd2009/small-churn/kdd_valid.csv
```


```
r211","Var212","Var213","Var214","Var215","Var216","Var217","Var218","Var219","Var220","Var221","Var222","Var223","Var224","Var225","Var22
6","Var227","Var228","Var229","Var230"]
  column_types:
["Enum","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric",
"Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric
","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numer
ic","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Num
eric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","N
umeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric",
"Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric
","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numer
ic","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Num
eric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","N
umeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric",
"Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric
","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numer
ic","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Numeric","Enum","Enum","Enum","En
um","Enum","Enum","Enum","Enum","Enum","Enum","Enum","Enum","Enum","Enum","Enum","Enum","Enum","Enum","Numeric","Enum","Enum","Enum","Enum
","Enum","Enum","Enum","Enum","Enum","Enum","Enum","Enum","Enum","Enum","Enum","Enum","Enum","Enum","Enum","Enum","Numeric"]
  delete_on_done: true
  check_header: 1
  chunk_size: 4194304
```


10.5s

Job

Run Time	00:00:07.78
Remaining Time	00:00:00.0
Type	Frame
Key	kdd_train.hex
Description	Parse
Status	DONE
Progress	100% Done.
Actions	View

7. To view the dataset, click on cache files generated by H2O, in this case “kdd_train.hex”. One can also check the basic stats in H2O: like number of rows, columns and compressed size of the file.

   Also one can check in table the max and means, cardinality etc. Also one can convert data type from the same.

kdd_train.hex

Actions: View Data | Split... | Build Model... | Predict | Download | Export

Rows	Columns	Compressed Size
39892	231	11MB

COLUMN SUMMARIES

label	type	Missing	Zeros	+Inf	-Inf	min	max	mean	sigma	cardinality	Actions
churn	enum	0	36969	0	0	0	1.0	0.0733	0.2606	2	Convert to numeric
Var1	int	39326	308	0	0	0	680.0	10.9965	41.2404	·	Convert to enum
Var2	int	38907	984	0	0	0	5.0	0.0051	0.1593	·	Convert to enum
Var3	int	38908	791	0	0	0	130668.0	466.5213	4735.9613	·	Convert to enum
Var4	int	38618	1256	0	0	0	27.0	0.1554	1.4184	·	Convert to enum
Var5	int	38721	712	0	0	0	6048550.0	247680.9095	670301.2058	·	Convert to enum
Var6	int	4426	787	0	0	0	114079.0	1321.8744	2541.9427	·	Convert to enum
Var7	int	4421	10415	0	0	0	35.0	6.8240	6.3084	·	Convert to enum
Var8	int	39892	0	0	0	·	·	0	-0	·	Convert to enum
Var9	int	39326	122	0	0	0	2300.0	48.4770	163.8910	·	Convert to enum
Var10	int	38721	722	0	0	0	12325590.0	360775.2989	897783.0925	·	Convert to enum
Var11	int	38908	0	0	0	8.0	40.0	8.5772	2.7229	·	Convert to enum
Var12	int	39448	240	0	0	0	1184.0	17.0270	70.8233	·	Convert to enum
Var13	int	4421	9782	0	0	0	197872.0	1260.9955	2878.9894	·	Convert to enum
Var14	int	38908	921	0	0	0	48.0	0.7175	3.8212	·	Convert to enum
Var15	int	39892	0	0	0	·	·	0	-0	·	Convert to enum
Var16	real	38721	45	0	0	0	434.9200	120.8566	73.2051	·	·
Var17	int	38618	759	0	0	0	1220.0	10.5495	47.0844	·	Convert to enum
Var18	int	38618	973	0	0	0	480.0	6.7630	26.2281	·	Convert to enum
Var19	int	38618	1246	0	0	0	27.0	0.2331	1.7135	·	Convert to enum

← Previous 20 Columns | → Next 20 Columns

CHUNK COMPRESSION SUMMARY

FRAME DISTRIBUTION SUMMARY

8. Once data is ready, we can build model

```
getModel “gbm-model”
```

CS

```
getModel "gbm-model"
```

Model

Model ID: gbm-model

Algorithm: Gradient Boosting Machine

Actions: Refresh | Predict... | Download POJO | Download Model Deployment Package (MOJO)

▸ MODEL PARAMETERS

▾ SCORING HISTORY - LOGLOSS