

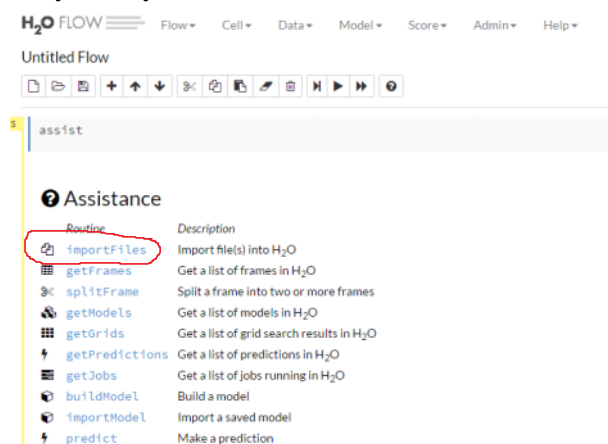
Implementing H2O using Web Flow

Wednesday, June 08, 2016 7:06 PM

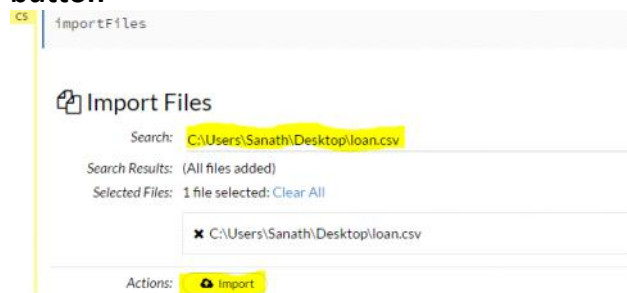
1. Start the command prompt to create the instance of **H2O**.
2. Set the path where you have downloaded the **H2O** file in command prompt
3. Execute the following command to create the instance of **H2O** : **java -jar h2o.jar**
4. You can view the instance of **H2O** in your browser by accessing the localhost (**localhost:54321**)
5. Import the loan.csv file in the **H2O** web flow to create the model

Snapshots

Step1: Import Files



Step2: Give the path for the loan.csv file in the Search tab, Add the selected file and then Hit the Import button



Step3: Once the path for the loan.csv file is selected, Hit the parse these files button



Step4: You can view the summary of the loan.csv file. Now change the data type of the bad_loan attribute from Numeric to Enum. Click the Parse button

Setup Parse

PARSE CONFIGURATION

Sources C:\Users\Sanath\Desktop\loan.csv

ID loan1.hex

Parser **CSV**

Separator **;**

Column Headers ☐ Auto

- ☒ First row contains column names
- ☐ First row contains data

Options ☐ Enable single quotes as a field quotation character

☒ Delete on done

EDIT COLUMN NAMES AND TYPES

Search by column name...

1	loan_amt	Numeric	5000	2500	2400	10000	5000	3000	5000	5375	6500		
2	term	Enum	36 months	60 months	36 months	36 months	36 months	36 months	60 months	60 months	60 months		
3	int_rate	Numeric	10.65	15.27	15.96	13.49	7.9	18.64	21.28	12.69	14.65		
4	emp_length	Numeric	10	0	10	10	3	9	4	0	5		
5	home_ownership	Enum	RENT	RENT	RENT	RENT	RENT	RENT	OWN	RENT	OWN		
6	annual_inc	Numeric	24000.0	30000.0	12252.0	49200.0	36000.0	40000.0	40000.0	15000.0	72000.0		
7	purpose	Enum	credit_card	car	small_business	other	wedding	car	small_business	other	debt_consolidation		
8	addr_state	Enum	AZ	GA	IL	CA	AZ	CA	CA	TX	AZ		
9	dti	Numeric	27.650000000000002	1.0	8.72	20.0	11.200000000000001	5.3500000000000005	5.55	18.080000000000002	16.12		
10	delinq_2yrs	Numeric	0	0	0	0	0	0	0	0	0		
11	revol_util	Numeric	83.7	9.4	98.5	21.0	28.3	87.5	32.6	36.5	20.6		
12	total_acc	Numeric	9	4	10	37	12	4	13	3	23		
13	bad_loan	Enum	0	1	0	0	0	0	1	1	0		
14	longest_credit_length	Numeric	26	12	10	15	7	4	7	7	13		
15	verification_status	Enum	verified	verified	not verified	verified	verified	verified	verified	verified	not verified		

Parse

Step5 : Check the job summary. Once it is completed you can build your model by hitting the Actions tab

Job

Run Time 00:00:00.383

Remaining Time 00:00:00.0

Type Frame

Key loan1.hex

Description Parse

Status DONE

Progress 100%

Done.

Actions View

Step6 : Now split the dataset into training and test set.

loan1.hex

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

Rows 163907

Columns 15

Compressed Size 6MB

COLUMN SUMMARIES

Table	type	Missing	Zero	-2of	-3of	Min	Q1	Median	Q3	sigma	cardinality	Actions
loan_amt	int	0	0	0	0	500.0	13074.1593	7993.5562	-	-	-	Convert to numeric
term	enum	0	129990	0	0	0	1.0	0.2876	0.4098	-	2	Convert to numeric
int_rate	real	0	0	0	0	5.4200	20.0000	13.7129	4.9919	-	-	Convert to numeric
emp_length	int	5864	14240	0	0	10.0	5.4044	2.4307	-	-	-	Convert to numeric
home_ownership	enum	0	1	0	0	0	5.0	-	-	-	6	Convert to numeric
annual_inc	real	4	0	0	0	1000.0	7143776.0	72315.6705	55070.9157	-	-	Convert to numeric
purpose	enum	0	2842	0	0	0	23.0	-	-	-	14	Convert to numeric
addr_state	enum	0	415	0	0	0	49.0	-	-	-	90	Convert to numeric
dti	real	0	270	0	0	0	39.9900	15.8815	7.5877	-	-	Convert to numeric
delinq_2yrs	int	20	139450	0	0	0	20.0	0.2374	0.6042	-	-	Convert to numeric
revol_util	real	193	13602	0	0	0	350.7000	94.0792	29.2894	-	-	Convert to numeric
total_acc	int	29	0	0	0	1.0	118.0	24.5797	11.6852	-	-	Convert to numeric
bad_loan	enum	0	133871	0	0	0	1.0	0.1830	0.2867	-	2	Convert to numeric
longest_credit_length	int	29	11	0	0	0	65.0	14.8543	6.9477	-	-	Convert to numeric
verification_status	enum	0	59155	0	0	0	1.0	0.6395	0.4802	-	2	Convert to numeric

CHUNK COMPRESSION SUMMARY

FRAME DISTRIBUTION SUMMARY

Step7 : Split the dataset into training(70%) and test(30%). Rename them as shown in snapshot and hit Create

Split Frame

Frame: loan1.hex

Splits: Ratio

0.70

0.30

Key

training

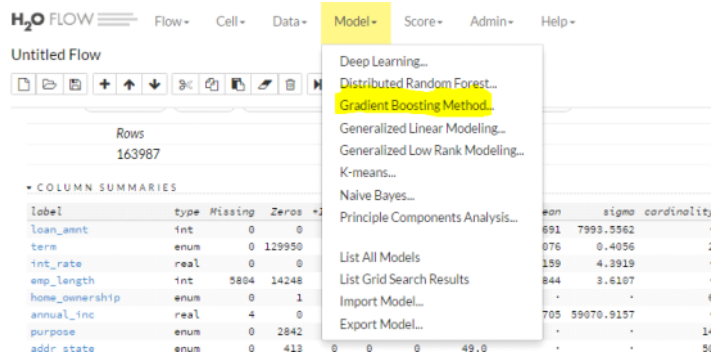
test

Add a new split

Seed: 597698

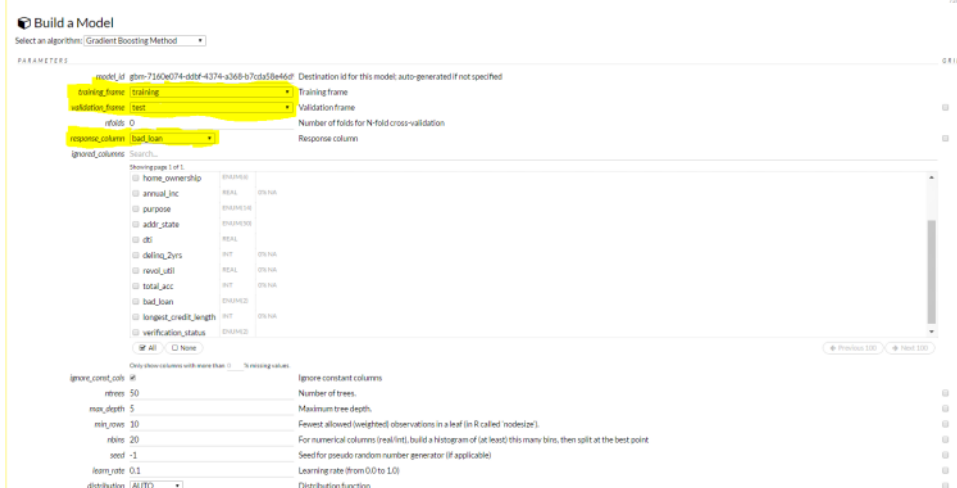
Create

Step8 : Create your model by selecting Model from Menu Bar. Select Gradient Boosting Method from the submenu Model

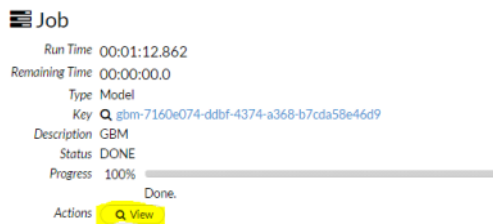


Step9 : Build the model. Make the following changes

- i. Select training as the training frame
- ii. Select test as the test frame
- iii. Select the response column as bad_loans
- iv. Check the score iteration checkbox
- v. Build the model by selecting the Build Model button at the end



Step 10 : Select the View action to view the results for your model



-----END-----

