Using GLM in H_2O Statistical Runtime for BigData

This is a brief description of how to use H_2O Statistical Runtime to ingest raw data and inspect the data for missing rows. We then use Generalized Linear Modeling (GLM) to identify the most important features for predicting conversions.

Introduction

Generalized Linear Modeling is the most powerful linear modeling toolkit at the hands of the modern data modeler. We present a simple REST-API for using the GLM on your datasets. This document describes using GLM via REST and JSON APIs.

We use the ADMM (Alternating Directions Method of Multipliers) technique for distributed version.

A simple web gui for GLM is also in the works and will be described in subsequent version.

Import dataset

1. **Import** the dataset via, ImportFolder or clicking store view (incase of HDFS launch)

Specify a folder whose files should be imported as keys to H2O. Please note that the folder must be local to all nodes and the path needs to be absolute.

| /home/hduser/datasets | Import Folder |

replication (optional)   import files recursively

Alternatively you can specify a URL to import from provided that the node you are connected to can reach it:

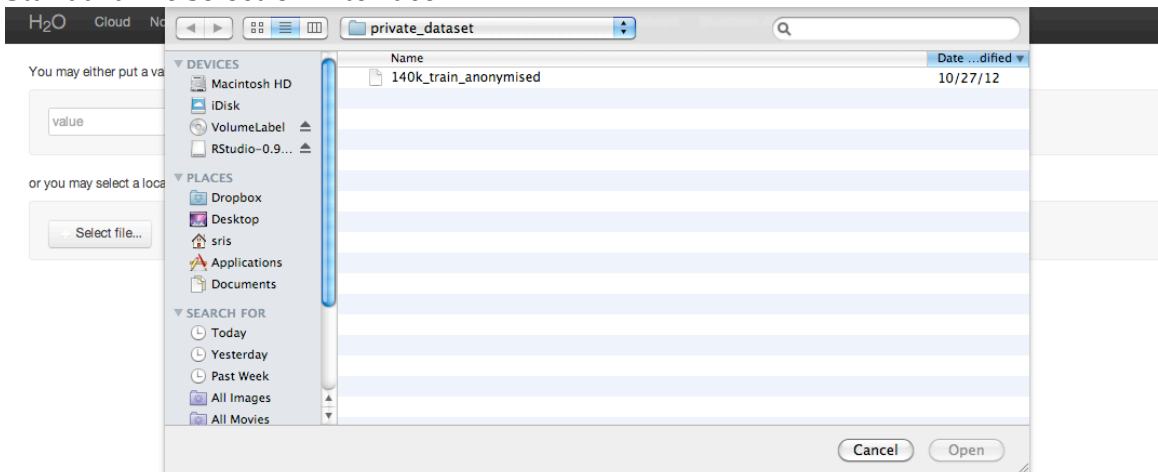| url | key (optional) | replication (optional) | Import URL |

2. **Put** operation can also be used for importing the dataset:

H₂O Cloud Node Get Put Timeline Import RF Debug View Progress View Network Shutdown All

You may either put a value:

| value | key (optional) | replication (optional) | Put |

or you may select a local file to be uploaded:

Select file...

Standard file selection interface -

H₂O Cloud No...   private_dataset

▼ DEVICES
Macintosh HD
iDisk
VolumeLabel ⏏
RStudio-0.9... ⏏

▼ PLACES
Dropbox
Desktop
sris
Applications
Documents

▼ SEARCH FOR
Today
Yesterday
Past Week
All Images
All Movies

Name   Date ...dified ▼
140k_train_anonymised   10/27/12

Cancel   Open

0xdata, Inc.
H₂O Statistical Runtime for BigData, GLM v1.0

3. **Parse** he dataset is a simple click through the link of the Key after Put or via the Store.

   This takes a few minutes for large datasets.

Parsed into //Users/sris/_work/hexbase_xem/smalldata/hhp_9_17_12.predict.data.hex in 762 msec

4. **Inspect** the data for a summary:

Generated from nfs://Users/sris/_work/hexbase_xem/smalldata/hhp_9_17_12.predict.data.gz_UNZIPPED by 'basic_parse'

224 Bytes-per-row * 70942 Rows = Totalsize 15891008
Parsed 217 columns

Column	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
Record offset	+0	+4	+5	+6	+7	+8	+9	+10	+11	+12	+13	+14	+15	+16	+17	+18	+19	+20	+21	+22	+23	+24	+25	+26	+27
Column bytes	4b	1b	1b	1b	1b	1b	1b	1b	1b	1b	1b	1b	1b	1b	1b	1b	1b	1b	1b	1b	1b	1b	1b	1b	1b
Internal scaling	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)
Min/Max	210 - 99997488	0 - 0	0 - 1	0 - 1	0 - 1	0 - 1	0 - 1	0 - 1	0 - 1	0 - 1	0 - 1	0 - 0	0 - 1	0 - 1	0 - 0	0 - 44	0 - 44	0 - 42	0 - 35	0 - 24	0 - 29	0 - 42	0 - 39	0 - 38	0 - 3!
μ	49915295.8566		0.1374	0.1543	0.1205	0.1279	0.0991	0.0839	0.0885	0.0614	0.0672	0	0.451	0.3648	0	3.3316	3.0321	2.1727	1.0083	0.9992	0.5838	0.4783	0.4041	0.2745	0.174
σ	28933883.8684		0.3443	0.3612	0.3256	0.334	0.2988	0.2772	0.284	0.24	0.2503	0	0.4976	0.4814	0	5.0641	4.8401	3.268	2.2944	1.6401	1.3533	1.9259	1.5787	1.6654	0.884
Rows																									
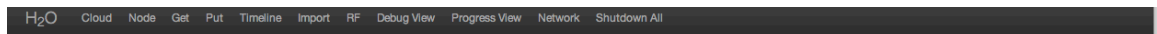
anony.hex
Generated from anony by 'basic_parse'

179 bytes-per-row * 142982 Rows = Totalsize 24.4 MB
Parsed 71 columns

Column	Month	sex	Day_Week	TimeofDay	WebApp	age	AnsweredSurvey	Srvy_Plan2DD	Srvy_bythngs_online	Has_bnk_AC	RegisteredOnline	Population	HouseholdsPerZipCode	WhitePopulation
Record offset	+0	+1	+2	+3	+4	+5	+6	+7	+8	+9	+10	+11	+15	+17
Column bytes	1b	1b	1b	1b	1b	1b	1b	1b	1b	1b	1b	4b	2b	4b
Internal scaling	(X+8)	(X)	(X+1)	(X)	(X)	(X+13)	(X)	(X)	(X)	(X)	(X)	(X+7)	(X)	(X)
Min/Max	8 - 11	0 - 2	1 - 7	0 - 5	0 - 1	13 - 95	0 - 1	0 - 1	0 - 1	0 - 0	0 - 1	7 - 114124	0 - 48391	0 - 86186
μ	9.4955		3.9189		0.7001	32.538	0.523	0.2859	0.0581	0	0.4126	31154.0178	12307.4286	15604.0944
σ	1.1164	0	1.8263	0	0.4582	11.2499	0.4995	0.4518	0.2339	0	0.4923	18872.1362	6873.2418	11667.8583
Rows missing data		759												
Row 0	10	M	5	16 to 18	1	47	1	0	0	0	1	21453	9825	20048
Row 1	10	F	5	16 to 18	1	43	1	0	0	0	0	14535	6384	5855
Row 2	10	F	5	16 to 18	1	21	1	0	0	0	1	23470	9469	12402

GLM

Generalized Linear Models are a powerful toolkit in any datamodeler's hands.
We invoke GLM using REST-API call –
Using L1 Regularization and a lambda that is high – 0.01 – we are able to detect features that donot matter.

```
http://localhost:54321/GLM?Key=anony.hex&Y=Converted&norm=L1&lambda=1
e-2&family=binomial&xval=10
```

| H₂O | Cloud | Node | Get | Put | Timeline | Import | RF | Debug View | Progress View | Network | Shutdown All |

GLM Parameters
family: binomial **link:** logit **norm:** L1 λ: 0.01 ρ: 0.01 α: 1.0
Coefficients

Month	sex	Day_Week	TimeofDay	WebApp	age	AnsweredSurvey	Srvy_Plan2DD	Srvy_bythngs_online	Has_bnk_AC	RegisteredOnline	Population	HouseholdsPerZipCode	WhitePopulation	BlackPopul
0	0	0	-0.0298	-0.2517	0.0253	0	0.0189	0	0	0.6597	0	0	0	0

Model SRC
y = 1/(1 + Math.exp(0.0298*x[TimeofDay] + 0.2517*x[WebApp] − 0.0253*x[age] − 0.0189*x[Srvy_Plan2DD] − 0.6597*x[RegisteredOnline] + 0.072*x[CityType] + 0.0216*x[division] + 0.0174*x[region] − 0.0123*x[CBSAPop2003] − 0.0107*x[Innovis_pass] + 0.0211*x[checkpointscore] + 0.0249*x[grade] + 0.059*x[white_percent] + 0.5789))

Validation

Degrees of freedom:	141812 total (i.e. Null); 141742 Residual
Null Deviance	55990236.128
Residual Deviance	155027.292
AIC	155169.292
Training Error Rate Avg	0.2639
False Positives	0.0063
False Negative	0.2576

10 fold Cross Validation
decision threshold = %threshold

	Mean	Variance
Error rate	0.2639	
True Positive	0.6794	0.0307
True Negative	0.0106	0.0001
False Negative	0.2444	0.004
False Positive	0.0061	

Individual Models
Model 1

	Y_real=0	Y_real=1
Y_model=0	10065	3722
	86	300

Table1: GLM with L1 Regularization produces a list of features that might not matter for predicting converted

Model SRC

```
y = 1/(1 + Math.exp(0.0298*x[TimeofDay] + 0.2517*x[WebApp] - 0.0253*x[age] - 0.0189*x[Srvy_Plan2DD] - 0.6597*x[RegisteredOnline] + 0.072*x[CityType] + 0.0216*x[division] + 0.0174*x[region] - 0.0123*x[CBSAPop2003] - 0.0107*x[Innovis_pass] + 0.0211*x[checkpointscore] + 0.0249*x[grade] + 0.059*x[white_percent] + 0.5789))
```

0xdata, Inc.
H₂O Statistical Runtime for BigData, GLM v1.0

Regression REST-API

Logistic Regression can be invoked on h2o via simple REST API. Here are a few simple examples:

family=binomial, LASSO

```
http://localhost:54321/GLM?Key=140k_train_anonymised.hex&Y=Converted&norm=L1&family=bin
omial
```

GLM on data 140k_train_anonymised.hex computed in 736[ms].

GLM Parameters
family: binomial **link:** logit **norm:** L1 **λ:** 0.1 **p:** 0.01 **α:** 1.0 **decision threshold:** 0.5

Coefficients

Month	sex	Day_Week	TimeofDay	WebApp	age	AnsweredSurvey	Srvy_Plan2DD	Srvy_bythngs_online	Has_bnk_AC	RegisteredOnline	Population	House
0	0	0	-0.0332	0	0	0	0	0	0	0.0953	0	0

Model SRC

$y = 1/(1 + Math.exp(0.0332*x[TimeofDay] - 0.0953*x[RegisteredOnline] + 0.0998*x[CityType] + 0.0766*x[division] + 0.0387*x[grade]))$

Validation

Degrees of freedom:	141812 total (i.e. Null); 141742 Residual
Null Deviance	55990236.128
Residual Deviance	168457.8537
AIC	168599.8537
Training Error Rate Avg	0.2789
False Positives	0.0028
False Negative	0.2761

family=gaussian, linear regression

```
http://localhost:54321/GLM?Key=covtype.20k.hex&Y=12&family=gaussian&X=1:8
```

GLM on data covtype.20k.hex computed in 5[ms].

GLM Parameters
family: gaussian **link:** identity **norm:** NONE **decision threshold:** 0.5

Coefficients
1 = -0.0001 2 = 0.0136 3 = 0.0007 4 = -0.0004 5 = -0 6 = 0.0096 7 = -0.0036 8 = 0.0083 Intercept = -2.2993

Model SRC

$y = -0.0001*x[1] + 0.0136*x[2] + 0.0007*x[3] - 0.0004*x[4] - 0*x[5] + 0.0096*x[6] - 0.0036*x[7] + 0.0083*x[8] - 2.2993$

Validation

Degrees of freedom:	19999 total (i.e. Null); 19991 Residual
Null Deviance	6349
Residual Deviance	3888.4499
AIC	3906.4499
Training Error Rate Avg	3888.4499

family=gaussian, linear regression

http://localhost:54321/GLM?Key=140k_train_anonymised.hex&Y=Converted&<mark>family=gaussian</mark>

GLM on data 140k_train_anonymised.hex computed in 123[ms].
Warnings: Failed to compute without normalization due to singular gram matrix. Rerun with L2 regularization and lambda = 1e-5

GLM Parameters
family: gaussian **link:** identity **norm:** NONE **decision threshold:** 0.5

Coefficients

Month	sex	Day_Week	TimeofDay	WebApp	age	AnsweredSurvey	Srvy_Plan2DD	Srvy_bythngs_online	Has_bnk_AC	RegisteredOnline	Population
0.0107	-0.0121	0.0003	-0.0006	-0.1261	0.0012	-0.02	0.0491	-0.0285	0	0.2937	0

Model SRC
```
y = 0.0107*x[Month] - 0.0121*x[sex] + 0.0003*x[Day_Week] - 0.0006*x[TimeofDay] - 0.1261*x[WebApp] + 0.0012*x[age] -
0.02*x[AnsweredSurvey] + 0.0491*x[Srvy_Plan2DD] - 0.0285*x[Srvy_bythngs_online] + 0.2937*x[RegisteredOnline] + 0*x[Population] -
0*x[HouseholdsPerZipCode] - 0*x[WhitePopulation] - 0*x[BlackPopulation] - 0*x[HispanicPopulation] - 0*x[AsianPopulation] -
0*x[HawaiianPopulation] - 0*x[IndianPopulation] - 0*x[OtherPopulation] + 0*x[MalePopulation] + 0*x[FemalePopulation] -
0.0262*x[PersonsPerHousehold] + 0*x[AverageHouseValue] + 0*x[IncomePerHousehold] - 0.016*x[MedianAge] + 0.0088*x[MedianAgeMale] +
0.0067*x[MedianAgeFemale] - 0*x[Elevation] - 0.0009*x[CityType] + 0.0007*x[division] - 0.0039*x[region] - 0.0014*x[TimeZone] +
0.008*x[DayLightSaving] + 0*x[NumberOfBusinesses] + 0*x[NumberOfEmployees] + 0*x[BusinessFirstQuarterPayroll] -
0*x[BusinessAnnualPayroll] + 0.0003*x[GrowthRank] + 0*x[GrowthHousingUnits2003] - 0*x[GrowthHousingUnits2004] +
0*x[GrowthIncreaseNumber] - 0.0022*x[GrowthIncreasePercentage] + 0*x[CBSAPop2003] - 0*x[CBSADivPop2003] + 0*x[DeliveryResidential] -
0*x[DeliveryBusiness] + 0*x[DeliveryTotal] - 0*x[PopulationEstimate] + 0*x[LandArea] + 0.0011*x[WaterArea] - 0*x[id] +
0.0159*x[Experian_pass] + 0.0203*x[Innovis_pass] - 0.0276*x[TU_pass] - 0.0173*x[LN_pass] + 0.0195*x[Experian_Cx] +
0.0345*x[Innovis_Cx] + 0.0007*x[TU_Cx] - 0.0109*x[LN_Cx] - 0.0001*x[checkpointscore] - 0.0021*x[levelonedecisioncode] -
0.0006*x[grade] + 0.1457*x[white_percent] - 0.0427*x[black_percent] - 0.1054*x[hispanic_percent] + 0.1472*x[male_percent] +
0.3733*x[female_percent] + 0.5205
```

Validation

Degrees of freedom:	141812 total (i.e. Null); 141742 Residual

As prompted we now apply L2 regularization and lambda=1e-5 to get –

L2 regularization, Linear Regression, lambda=1e-5

http://localhost:54321/GLM?Key=140k_train_anonymised.hex&Y=Converted&family=gaussian<mark>&norm=L2&lambda=1e-5</mark>

GLM Parameters

family: gaussian **link:** identity **norm:** L2 **λ:** 1.0E-5 **decision threshold:** 0.5

Coefficients

Month	sex	Day_Week	TimeofDay	WebApp	age	AnsweredSurvey	Srvy_Plan2DD	Srvy_bythngs_online	Has_bnk_AC	RegisteredOnline	Population
0.0152	-0.0121	0.0005	-0.0006	-0.0578	0.0139	-0.01	0.0222	-0.0067	0	0.1446	0.0841

Model SRC

y = 0.0152*x[Month] − 0.0121*x[sex] + 0.0005*x[Day_Week] − 0.0006*x[TimeofDay] − 0.0578*x[WebApp] + 0.0139*x[age] −
0.01*x[AnsweredSurvey] + 0.0222*x[Srvy_Plan2DD] − 0.0067*x[Srvy_bythngs_online] + 0.1446*x[RegisteredOnline] + 0.0841*x[Population] −
0.0285*x[HouseholdsPerZipCode] − 0.1389*x[WhitePopulation] − 0.1643*x[BlackPopulation] − 0.0182*x[HispanicPopulation] −
0.0336*x[AsianPopulation] − 0.0089*x[HawaiianPopulation] − 0.0081*x[IndianPopulation] − 0.0534*x[OtherPopulation] +
0.0996*x[MalePopulation] + 0.0693*x[FemalePopulation] − 0.0103*x[PersonsPerHousehold] + 0.0031*x[AverageHouseValue] +
0.0115*x[IncomePerHousehold] − 0.0703*x[MedianAge] + 0.0384*x[MedianAgeMale] + 0.0307*x[MedianAgeFemale] − 0.0032*x[Elevation] −
0.0009*x[CityType] + 0.0007*x[division] − 0.0039*x[region] − 0.0015*x[TimeZone] + 0.0079*x[DayLightSaving] +
0.0075*x[NumberOfBusinesses] + 0.0007*x[NumberOfEmployees] + 0.0046*x[BusinessFirstQuarterPayroll] − 0.0076*x[BusinessAnnualPayroll] +
0.0032*x[GrowthRank] − 0.0128*x[GrowthHousingUnits2003] − 0.0151*x[GrowthHousingUnits2004] + 0.0335*x[GrowthIncreaseNumber] −
0.0024*x[GrowthIncreasePercentage] + 0.0135*x[CBSAPop2003] − 0.0031*x[CBSADivPop2003] + 0.002*x[DeliveryResidential] −
0.0071*x[DeliveryBusiness] + 0.009*x[DeliveryTotal] − 0.0087*x[PopulationEstimate] + 0.0021*x[LandArea] + 0.0031*x[WaterArea] −
0.0137*x[id] + 0.0078*x[Experian_pass] + 0.0072*x[Innovis_pass] − 0.0061*x[TU_pass] − 0.0062*x[LN_pass] + 0.0063*x[Experian_Cx] +
0.0099*x[Innovis_Cx] + 0.0001*x[TU_Cx] − 0.0031*x[LN_Cx] − 0.0122*x[checkpointscore] − 0.0021*x[levelonedecisioncode] −
0.0007*x[grade] − 0.0441*x[white_percent] − 0.015*x[black_percent] − 0.0133*x[hispanic_percent] − 0.0029*x[male_percent] +
0.0029*x[female_percent] + 0.2851

Validation

Degrees of freedom:	141812 total (i.e. Null); 141742 Residual

10-fold cross validation

http://localhost:54321/GLM.json?Key=140k_train_anonymised.hex&Y=Converted&norm=L1&family=binomial&xval=10

H2O   Cloud   Node   Get   Put   Timeline   Import   RF   R   Debug View   Progress View   Network   Shutdown All

GLM Parameters

family: binomial **link:** logit **norm:** L1 **λ:** 0.1 **ρ:** 0.01 **α:** 1.0 **decision threshold:** 0.5

Coefficients

Month	sex	Day_Week	TimeofDay	WebApp	age	AnsweredSurvey	Srvy_Plan2DD	Srvy_bythngs_online	Has_bnk_AC	RegisteredOnline	Population	H
0	0	0	-0.0332	0	0	0	0	0	0	0.0953	0	0

Model SRC

y = 1/(1 + Math.exp(0.0332*x[TimeofDay] − 0.0953*x[RegisteredOnline] + 0.0998*x[CityType] + 0.0766*x[division] + 0.0387*x[grade]))

Validation

Degrees of freedom:	141812 total (i.e. Null); 141742 Residual
Null Deviance	55990236.128
Residual Deviance	168457.8537
AIC	168599.8537
Training Error Rate Avg	0.2789
False Positives	0.0028
False Negative	0.2761

10 fold Cross Validation

decision threshold = %threshold

	Mean	Variance
Error rate	0.2789	

JSON

The same commands can be used with GLM.json?Key=data.key&Y=Converted to get the JSON output of the model.

```
http://localhost:54321/GLM.json?Key=140k_train_anonymised.hex&Y=Converted&norm=L1&family=binomial&xval=10
```

{"key":"140k_train_anonymised.hex","h2o":"/192.168.1.50:54321","name":"Logistic regression","glmParams":
{"link":"logit","family":"binomial","threshold":0.5},"lsmParams":
{"norm":"L1","lambda":0.1,"rho":0.01,"alpha":1.0},"rows":142982,"time":739,"coefficients":
{"Month":0.0,"sex":0.0,"Day_Week":0.0,"TimeofDay":-0.03318011372765106,"WebApp":0.0,"age":0.0,"AnsweredSurvey":0.0,"Srvy_Plan2DD":0.0,"Srvy_bythngs
_online":0.0,"Has_bnk_AC":0.0,"RegisteredOnline":0.09529874820971607,"Population":0.0,"HouseholdsPerZipCode":0.0,"WhitePopulation":0.0,"BlackPopula
tion":0.0,"HispanicPopulation":0.0,"AsianPopulation":0.0,"HawaiianPopulation":0.0,"IndianPopulation":0.0,"OtherPopulation":0.0,"MalePopulation":0.0
,"FemalePopulation":0.0,"PersonsPerHousehold":0.0,"AverageHouseValue":0.0,"IncomePerHousehold":0.0,"MedianAge":0.0,"MedianAgeMale":0.0,"MedianAgeFe
male":0.0,"Elevation":0.0,"CityType":-0.09979244551563937,"division":-0.07660020861792383,"region":0.0,"TimeZone":0.0,"DayLightSaving":0.0,"NumberO
fBusinesses":0.0,"NumberOfEmployees":0.0,"BusinessFirstQuarterPayroll":0.0,"BusinessAnnualPayroll":0.0,"GrowthRank":0.0,"GrowthHousingUnits2003":0.
0,"GrowthHousingUnits2004":0.0,"GrowthIncreaseNumber":0.0,"GrowthIncreasePercentage":0.0,"CBSAPop2003":0.0,"CBSADivPop2003":0.0,"DeliveryResidentia
l":0.0,"DeliveryBusiness":0.0,"DeliveryTotal":0.0,"PopulationEstimate":0.0,"LandArea":0.0,"WaterArea":0.0,"id":0.0,"Experian_pass":0.0,"Innovis_pas
s":0.0,"TU_pass":0.0,"Choicepoint_pass":0.0,"LN_pass":0.0,"Experian_Cx":0.0,"Innovis_Cx":0.0,"TU_Cx":0.0,"Choicepoint_Cx":0.0,"LN_Cx":0.0,"checkpoi
ntscore":0.0,"levelonedecisioncode":0.0,"grade":-0.03865463488470411,"white_percent":0.0,"black_percent":0.0,"hispanic_percent":0.0,"male_percent":
0.0,"female_percent":0.0,"Intercept":0.0},"trainingSetValidation":
{"DegreesOfFreedom":141812,"ResidualDegreesOfFreedom":141742,"NullDeviance":"55990236.128","ResidualDeviance":"168457.8537","AIC":"168599.8537","tr
ainingSetErrorRate":"0.2789"},"trainingErrorDetails":
{"falsePositive":"0.0028","falseNegative":"0.2761","truePositive":"0.0026","trueNegative":"0.7185"},"xfactor":10,"threshold":0.5,"trueNegative":"0.
6828","trueNegativeVar":"0.0318","truePositive":"0.0013","truePositiveVar":"0","falseNegative":"0.2621","falseNegativeVar":"0.0043","falsePositive"
:"0.0027","models":[{"cm":[[10127,3965],[44,37]]},{"cm":[[10242,3851],[42,44]]},{"cm":[[10230,3869],[45,25]]},{"cm":[[10132,3967],[35,35]]},{"cm":
[[10217,3910],[40,36]]},{"cm":[[10263,3842],[43,42]]},{"cm":[[10181,3934],[45,32]]},{"cm":[[10067,4046],[34,37]]},{"cm":[[10250,3820],[36,46]]},
{"cm":[[10183,3948],[38,33]]}],"errRate":"0.2789"}

[1] Reference for the REST-API.

Key - (required) - key of the .hex data to run regression on

Y - (required) - column id / name of the response variable, if you run logistic regression it's values currently MUST be from {0,1}

X - (optional) - list of columns used as an input vector, either names or column indexes (or mixture of both), separated by ','.Ranges can be provided by supplying two values separated by ':',   e.g. X=1:3,5:10 is equivalent to X=1,2,3,5,6,7,8,9,10. Default is all columns except for Y.

-X - (optional) - negative X, these columns will be subtracted from the set of selected columns.   Thus, selected columns = X \ -X. Default is empty set.

family - (optional) - either "binomial" for logistic regression or "gaussian" for linear regression.Default is gaussian.

xval - (optional) - cross validation factor. supply 10 for 10fold cross validation. Default is 0 (no crossvalidation). Currently supported only for logistic regression.

threshold - (optional) - decision threshold used for prediction error rate computation. Only used by logistic regression. Default is 0.5.

norm - (optional) - norm for regularized regression. Possible values "L1", "L2", "NONE". Default is "NONE". If regularization is used, data will be REGULARIZED (so that it has zero mean and unit variance.)

lambda - (optional) - argument affecting the regularization, Default is 0.1.

rho - (optional) - used only with L1. Enhanced Lagrangian parameter. Default is 1.0.

alpha - (optional) - used only with L1. Over-relaxation parameter (typical values for alpha are between 1.0 and 1.8). Default is 1.0.

 [2] ADMM: Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers http://www.stanford.edu/~boyd/papers/admm_distr_stats.html