
H2O.ai

Case Study: Lending Club Dataset

	Column Name	Description	Unit	Role
1	loan_amnt	Requested loan amount	US dollars	Predictor
2	term	Longest term length	Months	Predictor
3	**int_term**	**Recommended interest rate**	**Rate**	**Response**
4	emp_length	Employment length	Years	Predictor
5	home_ownership	Housing status	Categorical	Predictor
6	annual_inc	Annual income	US dollars	Predictor
7	purpose	Purpose for the loan	Categorical	Predictor
8	addr_state	State of residence	Categorical	Predictor
9	dti	Debt to income ratio	Percent	Predictor
10	delinq_2yrs	Number of delinquencies in the past 2 years	Count	Predictor
11	revol_util	Revolving credit line utilized	Percent	Predictor
12	total_acc	Number of active accounts	Count	Predictor
13	**bad_loan**	**Bad loan indicator**	**Boolean**	**Response**
14	longest_credit_length	Age of oldest active account	Years	Predictor
15	verification_status	Income verification status	Boolean	Predictor

Case Study: Auto ML of Lending Club Dataset

```r
# Load package and connect to cluster
library(h2o)
h2o.init(max_mem_size = "6g")

# Import data and manage data types
train_path <- "https://raw.githubusercontent.com/h2oai/app-consumer-loan/master/data/loan.csv"
train <- h2o.importFile(train_path, destination_frame = "loan_train")
train["bad_loan"] = h2o.asfactor(train["bad_loan"])

# Set target and predictor variables
y <- "bad_loan"
x <- h2o.colnames(train)
x <- setdiff(x, c (y, "int_rate"))

# Use Auto ML to train models
aml <- h2o.automl(x = x, y = y, training_frame = train, max_runtime_secs = 300)
```

Case Study: Lending Club Dataset

	Column Name	Description	Unit	Role
1	loan_amnt	Requested loan amount	US dollars	Predictor
2	term	Longest term length	Months	Predictor
✖	**int_term**	**Recommended interest rate**	**Rate**	**Response**
4	emp_length	Employment length	Years	Predictor
5	home_ownership	Housing status	Categorical	Predictor
6	annual_inc	Annual income	US dollars	Predictor
7	purpose	Purpose for the loan	Categorical	Predictor
8	addr_state	State of residence	Categorical	Predictor
9	dti	Debt to income ratio	Percent	Predictor
10	delinq_2yrs	Number of delinquencies in the past 2 years	Count	Predictor
11	revol_util	Revolving credit line utilized	Percent	Predictor
12	total_acc	Number of active accounts	Count	Predictor
13	**bad_loan**	**Bad loan indicator**	**Boolean**	**Response**
14	longest_credit_length	Age of oldest active account	Years	Predictor