WSO2 Machine Learner 1.1.0 - Milestone Release 2

The second milestone release of WSO2 Machine Learner (ML) 1.1.0 contains deep learning algorithm support. Deep learning networks are very popular among data scientists due to their incredible ability to learn features from raw data. We are leveraging H2O's (http://h2o.ai/) Stacked Autoencoders Classifier.

For general information on WSO2 Machine Learner 1.1.0-m2 release, please visit our documentation https://docs.wso2.com/display/ML110/WSO2+Machine+Learner+Documentation

Stacked Autoencoders Algorithm

This is a multi-layer feed-forward artificial neural network that is trained with stochastic gradient descent using back-propagation. The network can contain a large number of hidden layers consisting of neurons with tanh, rectifier and maxout activation functions. Subsequent layers learns from activation from previous layers. Each compute node trains a copy of the global model parameters on its local data with multi-threading (asynchronously), and contributes periodically to the global model via model averaging across the network.

This is used as a classifier in ML.
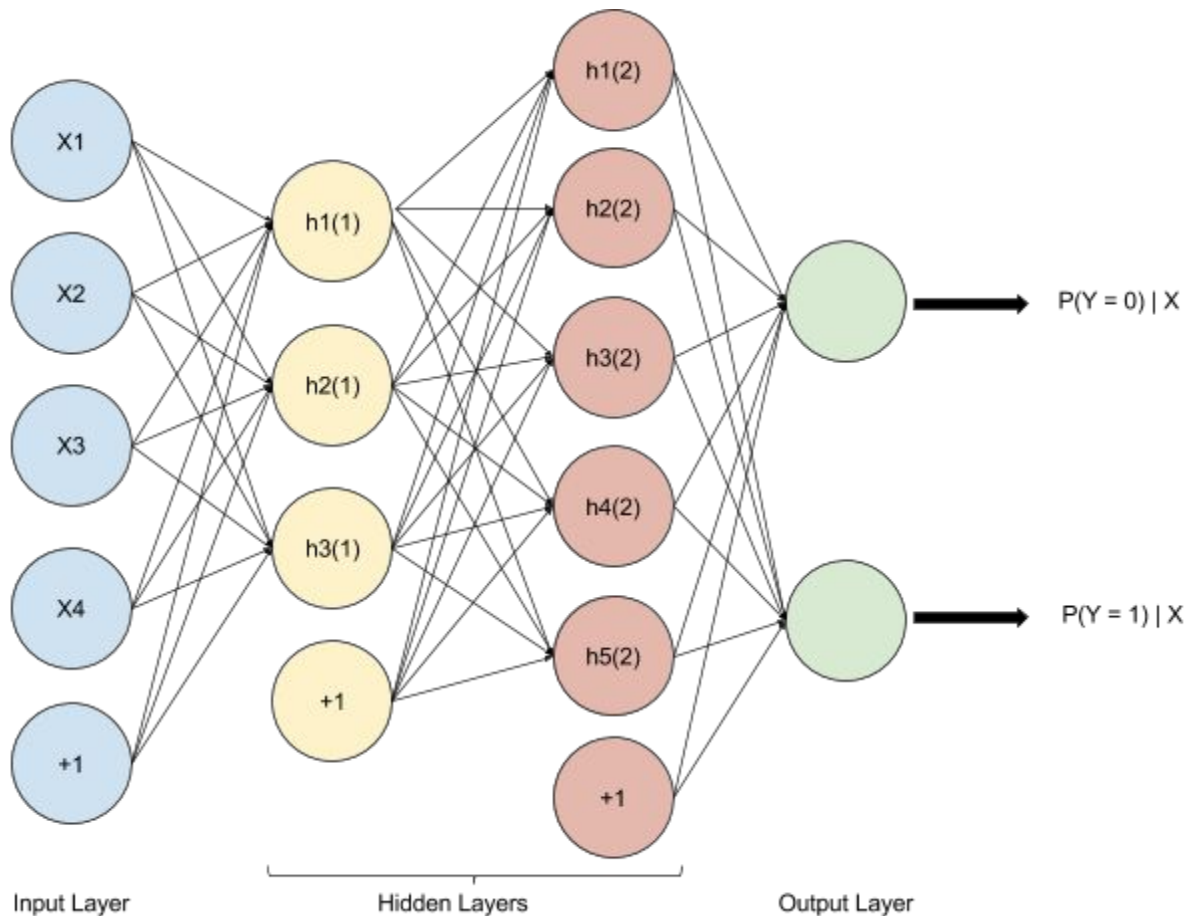
The above diagram shows a deep network with 4 inputs (4 features), 2 hidden layers and 2 outputs (2 classes to be predicted).

For more information on the implementation, please refer to the documentation: Implemented H2O Deeplearning and visualization for WSO2-ML.

Steps for Building a Deep Learning Model using WSO2 ML

Step 1 - Create an Analysis

Upload the dataset and create a new project.
As for the every model first you have to upload a dataset and create a new project. Then start a new analysis to build a deep learning model.

⌄ Diabetes_Project [created: 2015-11-26 08:57:08.51] ⚠ No analyses available   COMPARE MODELS 🗑 DELETE PROJECT

Diabetes project   Analysis name ❓

                                    wso2-ml-stacked-autoenco   ⊕ CREATE ANALYSIS

Step 2 - Algorithm Selection

In the Algorithm selection process there is a new category called Deep Learning. Select Stacked Autoencoders under that category.

Step 1	Step 2	Step 3	Step 4	Step 5
Preprocess	Explore	Algorithms	Parameters	Model

Algorithm

Algorithm name *

    STACKED AUTOENCODERS   ▾

Response variable *

    Class   ▾

Train data fraction *

    0.7

Step 3 - Hyper Parameters

In the parameter selection step you have to input necessary hyper parameters for the model.

- Batch Size - Number of training cases considered in each epoch

- Layer Sizes - Number of neurons in each layer. First layers is closest to the inputs and the last layer is furthest.
- Activation Type - The activation function (non-linearity) to be used for the neurons in the hidden layers. It can be one of the following:
  - Rectifier
  - RectifierWithDropout
  - Tanh
  - TanhWithDropout
  - Maxout
  - MaxoutWithDropout
- Epochs - Number of iterations the network is trained.

In addition to choosing hyperparameters, the deep network visualization will be show with the current configuration of parameters. Click "update visualization" if you want to change it after changing parameters.

Parameters

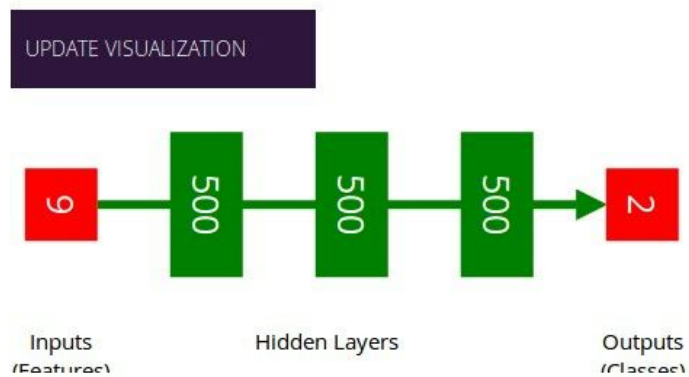Set Hyper-Parameters for Deeplearning\ **STACKED AUTOENCODERS**

Batch Size ❓

100

Layer Sizes ❓

500,500,500

Activation Type ❓

RectifierWithDropout

Epochs ❓

10

Deep Network Visualization

UPDATE VISUALIZATION

| Inputs (Features) | Hidden Layers | Outputs (Classes) |

Step 4 - Model Building

Then after selecting the dataset version you can build the model.

| Step 1	Step 2	Step 3	Step 4	Step 5
Preprocess	Explore	Algorithms	Parameters	Model

Model

Dataset version

diabetes_dataset-1.0.0

Step 5 - Model Summary

After successfully building the model you can view the model summary. In the summary you can get an overall idea about the performance of the build model. Model summary will provide measurements such as accuracy, confusion matrix and predicted vs. actual graph. You will be able to evaluate your model based on these metrics and may choose different parameters for a better model.

Model Summary [Accuracy: 61.16%]

Confusion Matrix

Total		Predicted	
		0	1
Actual	0	69	85
	1	9	79

Predicted vs. Actual

X-Axis: NumPregnancies

Y-Axis: PG2
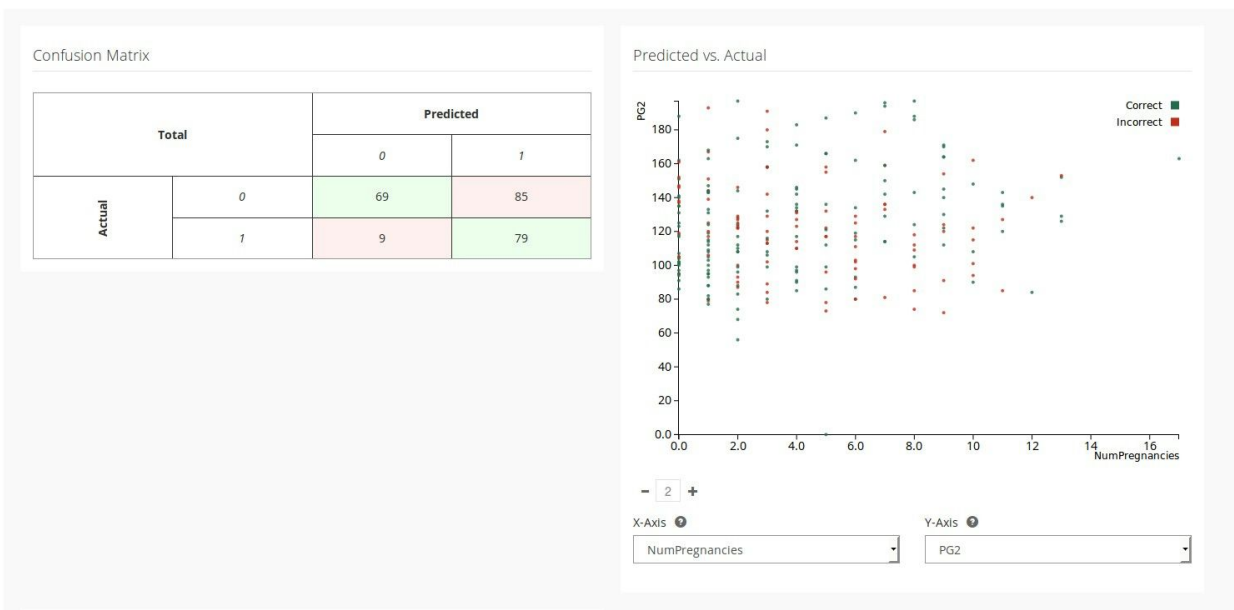
Moreover, you will be able to see the visualization of the deep network in the model summary.

Deep Network Visualization

8 — 500 — 500 — 500 → 2

Inputs   Hidden Layers   Outputs
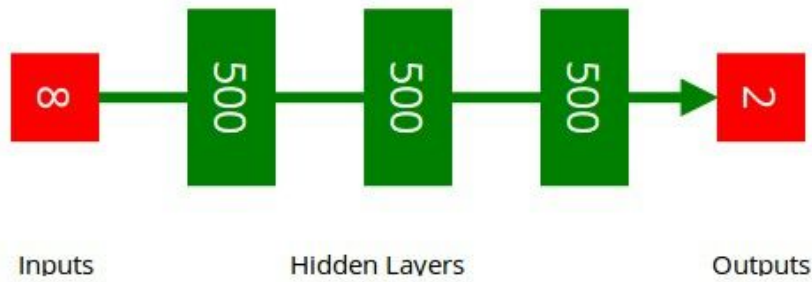
Step 6 - Prediction

This is where you can predict new data using built model. As the input, you have to give feature values of new data point or you can give new data as a batch using csv or tsv file. So after providing inputs to those values you will get the predictions for new data.

Predict

Prediction Source ❓

Feature values ▾

NumPregnancies *

2

PG2 *

3

DBP *

4

TSFT *

5

SI2 *

1

BMI *

4

DPF *

6

Age *

7

Predict

Deep Learning for WSO2 ML - Samples

Generating a Model Using the Stacked Autoencoders Algorithm

- Introduction
- Prerequisites
- Executing the sample
- Output of the sample

Introduction

This sample demonstrates how a model is generated out of a data set using the stacked autoencoders deep learning algorithm.

Prerequisites

Follow the steps below to set up the prerequisites before you start.

1. Download WSO2 Machine Learner, and start the server. For information on setting up and running WSO2 ML, see [Getting Started](#).
2. Download and install jq (CLI JSON processor). For instructions, see [jq Documentation](#).
3. If you are using Mac OS X, download and install GNU stream editor (sed). For instructions, see [GNU sed Documentation](#).

Executing the sample

Follow the steps below to execute the sample.

1. Navigate to <ML_HOME>/samples/default/stacked-autoencoders/ directory using the CLI.
2. Execute the following command to execute the sample: ./model-generation.sh

Output of the sample

Once the sample is successfully executed, you can view the prediction of the model. By default , the sample generates the model in the <ML_HOME>/models/ directory of your machine. For an example, the generated file is in the following format denoting the date and time when it was generated: wso2-ml-stacked-autoencoders-sample-analysis.Model.2015-11-26_12-00-46

Viewing the model

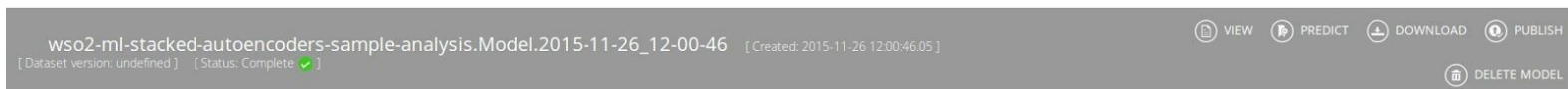You can view the summary of the built model using the ML UI as follows.

1. Log in to the ML UI from your Web browser using admin/admin credentials and the following URL: https://<ML_HOST>:<ML_PORT>/ml

2. Click the Projects button as shown below.

3. Click MODELS button of the new analysis which you created by executing the sample as shown below.

4. You view the built new model as shown below.

Viewing the model prediction

The sample executes the generated model on the
<ML_HOME>/samples/default/**stacked-autoencoders**/prediction-test data set, and it prints the value
[1.0] as the prediction result In the CLI logs.