Machine Learning Use Cases

with Azure ML
Agenda

* Use Cases
* Azure ML Studio
* Create a Model
* Operationalize Model
* Getting Started
Why Azure ML?

- Start in minutes
- No hardware or installation
- Only need a browser
- Use existing R or Python code
- One-click publishing
- Use what you need
Use Cases

* Data Cleaning and Preprocessing
* Build ML Models
* Operationalize your Model
* Technology Mashups
* Jupyter Notebooks
Microsoft Azure Machine Learning: Algorithm Cheat Sheet

This cheat sheet helps you choose the best Azure Machine Learning Studio algorithms for your predictive analytics solution. Your decision is driven by both the nature of your data and the question you’re trying to answer.

ANOMALY DETECTION
- One-class SVM
  - >100 features, aggressive boundary
- PCA-based anomaly detection
  - Fast training

CLUSTERING
- K-means
  - Finding unusual data points
  - Discovering structure

MULTI-CLASS CLASSIFICATION
- Fast training, linear model
  - Multiclass logistic regression
- Accuracy, long training times
  - Multiclass neural network
- Accuracy, fast training
  - Multiclass decision forest
- Accuracy, small memory footprint
  - Multiclass decision jungle
- Depends on the two-class classifier, see notes below
  - One-v-all multiclass

REGRESSION
- Ordinal regression
  - Data in rank ordered categories
- Poisson regression
  - Predicting event counts
- Fast forest quantile regression
  - Predicting a distribution
- Linear regression
  - Fast training, linear model
- Bayesian linear regression
  - Linear model, small data sets
- Neural network regression
  - Accuracy, long training time
- Decision forest regression
  - Accuracy, fast training
- Boosted decision tree regression
  - Accuracy, fast training, large memory footprint

TWO-CLASS CLASSIFICATION
- Two-class SVM
  - >100 features, linear model
- Two-class averaged perceptron
  - Fast training, linear model
- Two-class logistic regression
  - Fast training, linear model
- Two-class Bayes point machine
  - Fast training, linear model

Two-class decision forest
- Accuracy, fast training
- Two-class boosted decision tree
- Accuracy, large memory footprint
- Two-class decision jungle
- Two-class locally deep SVM
- Two-class neural network
Custom Algorithms

- **R**
  - Execute R - Data Processing & Cleanup
  - Create R Model - Predictions

- **Python**
  - Execute Python - Data Processing & Cleanup
Demo: ML Studio
Demo: R Model
Demo: Publish a Service
Getting Started

https://studio.azureml.net
# Free vs. Standard

**Pricing Details**

Machine Learning is offered in two tiers: Free and Standard. Features by tier are compared in the table below:

<table>
<thead>
<tr>
<th></th>
<th>FREE</th>
<th>STANDARD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authentication</td>
<td>Microsoft account (does not require an Azure subscription or a credit card)</td>
<td>Requires Azure subscription</td>
</tr>
<tr>
<td>Max Number of Modules per Experiment</td>
<td>100</td>
<td>Unlimited</td>
</tr>
<tr>
<td>Max Experiment Duration</td>
<td>1 hour</td>
<td>None</td>
</tr>
<tr>
<td>Max Storage Space</td>
<td>10 GB</td>
<td>Unlimited - BYO</td>
</tr>
<tr>
<td>Execution / Performance</td>
<td>Single node</td>
<td>Multiple nodes</td>
</tr>
<tr>
<td>Staging Web API</td>
<td>Yes (Throttled performance)</td>
<td>Yes (Selectable performance)</td>
</tr>
<tr>
<td>Production Web API</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>SLA</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Questions?

Contact Info:

mchenry19@gmail.com
@CAMCHENRY
http://cmchenry.com
http://www.linkedin.com/in/cmchenry
https://plus.google.com/+chrismchenry
Backups
## Pricing

### Standard tier pricing

<table>
<thead>
<tr>
<th>Service</th>
<th>Description</th>
<th>Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>ML Seat Subscription</td>
<td></td>
<td>Monthly Fee</td>
</tr>
<tr>
<td>ML Studio Usage</td>
<td></td>
<td>Hourly</td>
</tr>
<tr>
<td>ML API Usage</td>
<td></td>
<td>Hourly</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Transactions</td>
</tr>
</tbody>
</table>

Hourly charges only apply to active use of the service. Where multiple meters are present they are applied concurrently.
New Experiment
Regression Experiment
Visualize Data
Score a Model
Evaluate Models
# Web Services

<table>
<thead>
<tr>
<th>NAME</th>
<th>CREATED ON</th>
</tr>
</thead>
<tbody>
<tr>
<td>Automobile Price Prediction (Scoring Exp.)</td>
<td>6/2/2015 2:53:55 PM</td>
</tr>
<tr>
<td>Credit Risk - Scoring (Scoring Exp.)</td>
<td>4/19/2015 3:36:28 PM</td>
</tr>
</tbody>
</table>
Notebooks
Datasets

<table>
<thead>
<tr>
<th>NAME</th>
<th>SUBMITTED BY</th>
<th>DESCRIPTION</th>
<th>DATA TYPE</th>
<th>CREATED</th>
<th>SIZE</th>
</tr>
</thead>
<tbody>
<tr>
<td>AzureML-RPackages.csv</td>
<td>cmcherry</td>
<td></td>
<td>GenericCSV</td>
<td>9/19/2015 9:00:47 PM</td>
<td>64.49 KB</td>
</tr>
<tr>
<td>utilities.zip</td>
<td>cmcherry</td>
<td></td>
<td>Zip</td>
<td>6/12/2015 9:08:49 PM</td>
<td>1.96 KB</td>
</tr>
<tr>
<td>mshNews.csv</td>
<td>cmcherry</td>
<td></td>
<td>GenericCSV</td>
<td>6/2/2015 5:48:25 PM</td>
<td>96.69 KB</td>
</tr>
<tr>
<td>mshPrice.csv</td>
<td>cmcherry</td>
<td></td>
<td>GenericCSV</td>
<td>5/31/2015 3:58:22 PM</td>
<td>648.67 KB</td>
</tr>
<tr>
<td>german.csv</td>
<td>cmcherry</td>
<td>Downloaded from <a href="http://archive.ics.uci">http://archive.ics.uci</a>...</td>
<td>GenericCSV</td>
<td>4/7/2015 9:08:59 PM</td>
<td>77.92 KB</td>
</tr>
</tbody>
</table>
# Trained Models

<table>
<thead>
<tr>
<th>Name</th>
<th>Submitted By</th>
<th>Description</th>
<th>Data Type</th>
<th>Created</th>
</tr>
</thead>
<tbody>
<tr>
<td>Credit Risk - Scoring [trained model]</td>
<td>cmchenry</td>
<td></td>
<td>LearnerDotNet</td>
<td>4/18/2015 3:28:35 PM</td>
</tr>
</tbody>
</table>
Settings

NAME: AUTHORIZATION TOKENS

WORKSPACE NAME: Demo

WORKSPACE DESCRIPTION: Default workspace.

WORKSPACE TYPE: Free

WORKSPACE ID: (Blue box indicating unused space)

WORKSPACE STORAGE: 0.07 GB, 1% of 10 GIGABYTES

Want more storage? Get the standard version learn more
R Model

Naive Bayes Classifier R Model

Properties

Experiment Properties
END TIME: 9/23/2015 11:44:50 AM
STATUS CODE: Finished
STATUS DETAILS: None

Summary
This experiment demonstrates how to use the "Create R Model" module to train and score a model, and use "Execute Python Script" to evaluate a model using breast cancer classification as an example.

Description
Enter the detailed description for your experiment.

Quick Help
Split Module
R Training Model

Naive Bayes Classifier R Model

Create R Model

Use R Script to create Naive Bayes classifier

Train Model

Score Model

Execute Python Script

Properties

Trainer R script

1. # Input: dataset
2. # Output: model
3. # The code below is an example which can be replaced with your own code.
4. # See the help page of "Create R Model" module for the list of predefined functions.
5. library(e1071)
6. features <- get.feature.columns(dataset)
7. labels <- as.factor(get.label.column(dataset))
8. train.data <- data.frame(features, labels)
9. feature.names <- get.feature.column.names(dataset)
10. names(train.data) <- c(feature.names, "Class")
11. model <- naiveBayes(Class ~ ., train.data)

Scorer R script

1. # Input: model, dataset
2. # Output: scores
3. # The code below is an example which can be replaced with your own code.
4. # See the help page of "Create R Model" module for the list of predefined functions.
5. # Creates an R model using custom resources

Quick Help

(more help...)
R Scoring Model
Python Eval of R Model

```python
def azureml_main(dataframe):
    import matplotlib
    matplotlib.use("agg")
    from sklearn.metrics import accuracy_score, precision_score, recall_score, roc_auc_score, roc_curve
    import pandas as pd
    import numpy as np
    import matplotlib.pyplot as plt

    scores = dataframe[['Class', 'classes', 'probabilities']]
    ytrue = scores['Class']
    ypred = np.array([float(val) for val in scores['classes']])
    probabilities = scores['probabilities']

    accuracy, precision, recall, auc = \
        accuracy_score(ytrue, ypred), \
        precision_score(ytrue, ypred), \
        recall_score(ytrue, ypred), \
        roc_auc_score(ytrue, probabilities)

    metrics = pd.DataFrame();
    metrics["Metric"] = ["Accuracy", "Precision", "Recall", "AUC"];
    metrics["Value"] = [accuracy, precision, recall, auc]

    # Plot ROC Curve
    fpr, tpr, thresholds = roc_curve(ytrue, probabilities)
    fig = plt.figure()
    axis = fig.gca()
    axis.plot(fpr, tpr, linewidth=3)
    axis.grid("on")
    axis.set_xlabel("False positive rate")
```
R Model Evaluate
R Model ROC Curve
R Model Create Scoring

Naive Bayes Classifier R Model

- Breast cancer data
- Split
- Create R Model
  Use R Script to create Naive Bayes classifier
- Train Model
- Score Model
- Execute Python Script

Properties

Experiment Properties
START TIME: 9/23/2015 2:05:24 PM
END TIME: 9/23/2015 2:05:27 PM
STATUS CODE: Finished
STATUS DETAILS: None

Summary
This experiment demonstrates how to use the "Create R Model" module to train and score a model, and use "Execute Python Script" to evaluate a model using breast cancer classification as an example.

Description
Enter the detailed description for your experiment.
R Model Scoring Exp
R Model Scoring Results
R Model Publish
R Model Service
R Model Req/Resp

Request Response API Documentation for Naive Bayes Classifier R Model [Scoring Exp.]

Updated: 09/23/2015 20:10

No description provided for this web service.

- Previous version of this API
- Submit a request
- Input Parameters
- Output Parameters
- ASP.Net App Template for RPS
- Sample Code

OData Endpoint Address

Request

<table>
<thead>
<tr>
<th>Method</th>
<th>Request URI</th>
<th>HTTP Version</th>
</tr>
</thead>
</table>

Note: You may omit the details parameter from the query string. This would cause ColumnTypes to be omitted from the output

Request Headers

<table>
<thead>
<tr>
<th>Request Header</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authorization: Bearer abc123</td>
<td>Required. Pass the API Key here. Obtain this key from the publisher of the API.</td>
</tr>
<tr>
<td>Content-Length</td>
<td>Required. The length of the content body.</td>
</tr>
</tbody>
</table>
**R Model Sample Code**

<table>
<thead>
<tr>
<th>classes</th>
<th>Categorical</th>
<th>0, 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>probabilities</td>
<td>Numeric</td>
<td></td>
</tr>
</tbody>
</table>
Invoking R Model
Testing a Web Service

Microsoft Azure Machine Learning

naive bayes classifier r model [scoring exp.]

DASHBOARD

Configuration

General

Published experiment

View snapshot

View latest

Description

No description provided for this web service.

API key

Default Endpoint

REQUEST/RESPONSE

TEST

APPS

LAST UPDATED

9/23/2015 1:14:57 PM

9/23/2015 1:14:57 PM

2 OPERATIONS HAVE COMPLETED

- Naive Bayes Classifier R Model [Scoring Exp.] test returned ["0","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3","3",...