SAS Machine Learning and other Analytics: Trends and Roadmap

Sascha Schubert
Sberbank
8 Sep 2017
How „Big“ Analytics will Change Organizations

Optimization and Innovation
- Optimizing existing processes
  - Customer service
  - Pattern detection

Transformation and Disruption
- Data Monetarization
  - Definition of new business models
  - Collaboration with new industries

Both benefit from the same technology trends
“There were 5 Exabytes of information created between the dawn of civilization through 2003, but that much information is now created every 2 days.”

Eric Schmidt, Google CEO 2010
SAS® Visual Data Mining and Machine Learning

- K-means and K-modes Clustering
- Linear and Logistic Regression
- Generalized Linear Models
- Partial Least Squares Regression
- Nonlinear Regression
- Principal Component Analysis
- Robust PCA
- Moving Window PCA
- Quantile Regression
- Decision Trees*
- Factorization Machines*
- Gradient Boosting*
- Neural Networks*
- Random Forest*
- Support Vector Machines*
- Boolean Rules
- Text Mining
- Network Analytics/Community Detection
- Autotuning

* Supports Autotuning

**Discovery**

**Deployment**

**Data**

- Assess Supervised Models
- Creates score code
- Multi Threaded Data Step
- DS2
- SQL
- Variable Binning
- Variable Cardinality Analysis
- Sampling and Partitioning
- Missing Value Imputation
- Variable Selection
- Transpose
- Image processing

* Supports Autotuning
AUTOTUNING HYPERPARAMETERS

• Highly data dependent
• Related to model complexity

• Auto Tuning:
  • Automate hyperparameters search and find the optimal set
  • Maximize predictability on independent data set
  • Aims to avoid over-fitting by controlling model complexity
  • Creates more accurate models faster vs hand tuning
  • SAS auto tuning leverages world class SAS optimization engines
High Frequency Analytics Applications

• Many use cases are focused on anomaly detection
  • Fraud detection
  • Cyber-security intrusion detection
  • Process degradation (capitally intensive assets)
From Images to Data

Image

Preprocess and transform into table

<table>
<thead>
<tr>
<th>c1</th>
<th>c2</th>
<th>c3</th>
<th>c4</th>
<th>c5</th>
<th>c6</th>
<th>c7</th>
<th>c8</th>
<th>c9</th>
</tr>
</thead>
<tbody>
<tr>
<td>208.0</td>
<td>220.0</td>
<td>225.0</td>
<td>232.0</td>
<td>237.0</td>
<td>244.0</td>
<td>250.0</td>
<td>254.0</td>
<td>255.0</td>
</tr>
</tbody>
</table>

Analyze

Copyright © SAS Institute Inc. All rights reserved.
SAS® Viya™: New Open Architecture
Different Interfaces for Different User Personas

Interactive GUI

Programming

API Interfaces

SAS® Viya™
SAS® Visual Data Mining and Machine Learning

GUI Interface – SAS Visual Analytics

- Machine Learning
  - Forest
  - Gradient Boosting
  - Neural Networks
  - Support Vector Machines
  - Factorization Machines

- Statistics
  - Linear Regression
  - Logistic Regression
  - GLM Regression
  - Clustering (k-means)
  - Decision Tree
SAS® Visual Data Mining and Machine Learning
Programmatic Interface - SAS Studio

- Web Interface
- Interactive Program Editor
- Snippets
- Tasks
- Program Generator
SAS Viya Programming
As viewed through Legos...

SAS Studio can include one or more **procedures** to create a **task**.

Stack one or more **action sets** to create **procedures**

Stack one or more **actions** to create **action sets**

Each **action** has one or more parameter settings.

Actions are at the heart of every CAS procedure
SAS® Viya™ and Open Source

• SAS Scripting Wrapper for Analytics Transfer (SWAT) package is an API from Python or R to CAS.
• SWAT allows users to load data into memory and apply CAS actions to transform, summarize, model and score the data
• Integration of SAS Analytics in Python/R code
• Jupyter Notebook support
**SAS® Viya™: New Interface APIs**

**Different Languages – Same Power**

**SAS® Viya®**

Controller

Workers

---

**Translated Action**

```python
import sasviyavisual as viyavisual

# Sample R code
vf <- defCasTable(s, 'hmeq')
hd <- head(vf, 10)
```

```R
vf <- defCasTable(s, 'hmeq')
hd <- head(vf, 10)
```

---

```python
df = s.CASTable('hmeq')
df.head(10)
```

---

```R
df <- defCasTable(s, 'hmeq')
head(df, 10)
```

---

```SAS
proc print data = hmeq (obs = 10);
run;
```
SAS® ENTERPRISE MINER VIYA™ NODE

Increase the Reach

- Model comparison
- Testing new SAS Viya algorithms
- Simultaneous processing of all model trainings
- Model ensembles
- Integration with SAS Enterprise Miner Metadata
- Management, deployment und monitoring of SAS Viya models
Different Interfaces for Different User Personas

One Integrated Solution for Different User Types

- Business Analyst
- SAS Statistician
- SAS Data Scientist
- Open Source Data Scientist
- IT and Application Mngt.
Analytics Lifecycle

Integrated Machine Learning

Data Prep
- Analytical data preparation
- Batch Deployment

Model Building
- Segmented Models
- Full-scale process flow
- Open Source Models

Model Deployment
- Model Monitoring
- In-Database deployment
- Data driven Rules

Deployment Targets
- Batch
- In Database
- Hadoop
- Decision Service
- In Stream
Useful Websites
Developer.sas.com, Communities.sas.com
USE CASE - Real Time NBA

The customer makes a purchase in a store

Billing system stops the transaction as it exceeds expense limit

Event: transaction denied is considered relevant for real-time decision management system

RT Decision Engine

Event

Context information
- Customer profile
- Balance account
- Expense limit
- Purchase amount

Notice is sent on the Bank web portal, the bank APP, or via SMS

You have reached your credit card limit. We can increase your limit by 500 Euro. Please answer with Yes

NBA is calculated using: customer profile and eligibility to that specific offer

Instant expense limit increase (500€)
Automated Decisions

ChatBot

Hi, I'm Sir Botsalot

Nathan

****

Lets chat!
SAS Viya™ Virtual Learning

Support your adoption

- Free resources
  - Getting Started training
  - Video Libraries
  - e-Learning
- Topics Include
  - Administration
  - Programming & Analytics
  - SAS Visual Analytics
  - Open Source Integration

Get Started with SAS® Viya™ Enablement
Some highlights

- Custom pipelines
- Parallel execution
- SAS best practice pipelines (i.e. Rapid Predictive Modeler)
- Defining and sharing custom reusable nodes
- Ability to deploy models into databases directly
- Score code generation and score code APIs
- Import SAS 9.4 score code into pipeline comparison
- SAS Code integration
Deep Learning

SAS Vision

Provide market leading platform for the rapid and seamless extraction of knowledge from data of any size, type, or complexity.

Provide deep learning for SAS Solutions (Fraud, Customer Intelligence and Security Intelligence) and for customers to embed into their applications.

Deep Forward Networks
Convolutional Neural Networks

Recurrent Neural Networks
Autoencoders
Neural Networks

- Perceptron
- Feed Forward Networks (FF); Radial Basis Networks (RBF)
- Support Vector Machines (SVM)

Deep Learning

- Deep Neural Network (DNN)
- Convolutional Neural Networks (CNN)
- Recurrent Neural Networks (RNN)
- Auto Encoder (AE)

http://www.asimovinstitute.org/neural-network-zoo/
Artificial Intelligence (AI) is the science of enabling computers to perform tasks that typically require human intelligence to complete.
Thank you!

Sascha Schubert
sascha.schubert@sas.com