New SAS Platform
SAS Platform

What Do I Need?

• Build The CHAMPION Model
• Deliver Consistent Results
• Explain the Prediction Results
• Deploy the Model by One Click
• Fuel Business Needs with Governance
Build The CHAMPION Model

Increase Productivity Using Distributed In-Memory Processing

CAS (Cloud Analytic Services)
The “engine” of Viya Platform

CAS is a platform for high-performance analytics, in memory and distributed computing

You develop fast and use the entire data
Build The CHAMPION Model

Algorithms - Machine Learning and Deep Learning

**Statistics**
- Cox Proportional Hazards
- Decision Trees
- Design Matrix
- General Additive Models
- Generalized Linear Models
- K-means and K-modes Clustering
- Linear Regression
- Logistic Regression
- Nonlinear Regression
- Ordinary Least Squares Regression
- Partial Least Squares Regression
- Pearson Correlation
- Principal Component Analysis
- Quantile Regression
- Shewhart Control Chart Analysis

**Machine Learning**
- Bayesian Networks
- Boolean Rules
- Factorization Machines
- Frequent Item Set Mining
- Gradient Boosting
- K Nearest Neighbor
- Market Basket Analysis
- Moving Windows PCA
- Network Analytics/Community Detection
- Random Forest
- Robust PCA
- Support Vector Data Description
- Support Vector Machines
- Text Mining
- Variable Clustering

**Deep Learning**
- Deep Forward Neural Networks (DNNs)
- Convolutional Neural Networks (CNNs)
- Support VGG-like models
- Support ResNet models
- Recurrent Neural Networks (RNNs)
- Support LSTM model
- Support GRU model
- Autoencoders for neural networks
- GPU Support
- Image processing extensions
- 2D/3D medical image visualization

Latest and Greatest algorithms producing the same results across different languages!
Build The CHAMPION Model

Achieve the Champion Model- Automated Hyperparameter Tuning

Parallelized and Optimized Hyperparameter Tuning!
Deliver Consistent Results

Different Programming Languages

Programming Interfaces

SAS

```
proc print data = hmeq (obs = 10);
run;
```

Python

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

R

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

Translated Command

```
{table.fetch}
  table.name = "hmeq"
  from = 1 to = 10
```
Deliver Consistent Results

Consistent Results - for All Set of Skillsets

Visual Interfaces

Programming Interfaces

<table>
<thead>
<tr>
<th>Descri</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 Number of Observations Read</td>
<td>1788</td>
</tr>
<tr>
<td>1 Number of Observations Used</td>
<td>1788</td>
</tr>
<tr>
<td>2 Misclassification Error (%)</td>
<td>7.9418344519</td>
</tr>
</tbody>
</table>

Validation Misclassification: 0.0794
Deliver Consistent Results

Deliver Consistent Results Throughout Time

• Monitor the model and re-train it as soon as the performance drops

Visually Monitor the performance and set triggers for re-training!
## Explain The Prediction Results

### Interpretability Techniques- Overcoming the Challenge of Clear Reason

<table>
<thead>
<tr>
<th>Question</th>
<th>Technique</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>What are the top inputs?</td>
<td>Decision Tree Surrogate Model</td>
<td>Find the main drivers for an output</td>
</tr>
<tr>
<td>How do the drivers work?</td>
<td>Partial Dependence (PD)</td>
<td>Gives marginal effect of selected input variable (or multiple variables on target)</td>
</tr>
<tr>
<td></td>
<td>Individual Conditional Expectation (ICE)</td>
<td>Can help to identify subgroups (additive) effects and interactions</td>
</tr>
<tr>
<td>What is the explanation for a particular predication?</td>
<td>Local Interpretable Model-Agnostic Explanation (LIME)</td>
<td>Gives explanations for individuals prediction from a classifier</td>
</tr>
</tbody>
</table>

Interpret the predictions visually by a click!
Explain The Prediction Results

PD and ICE Explanation Plots
Explain The Prediction Results

LIME- Data Point Level Prediction Explanation

Local Interpretable Model-agnostic Explanations (LIME)

LIME builds an interpretable model of explanatory data samples at local areas in the analyzed data.

Local
A cluster centroid is chosen, along with other data points within close proximity.

Interpretable
Builds a linear regression model to fit the data

Model-agnostic
Builds a linear regression around local points, regardless of how the original prediction was generated.

"Why Should I Trust You?" - Explaining the Predictions of any Classifier
Marco Tulio Ribeiro, Sameer Singh, Carlos Guestrin, ACM SIGMOD, 2016
Deploy the Model By One Click

Different Deployment Options

- Champion
- Registered
- Name
  - Gradient Boosting
  - Gradient Boosting
  - Level Encoding + Gradient Boosting
  - Level Encoding + Gradient Boosting

Error Plot

- Average Squared Error

Variable Importance

<table>
<thead>
<tr>
<th>Variable</th>
<th>Train Im</th>
<th>Importance</th>
<th>Relative</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEBTINC</td>
<td>9.8679</td>
<td>20.2091</td>
<td>1</td>
</tr>
<tr>
<td>NINQ</td>
<td>3.8805</td>
<td>1.0022</td>
<td>0.3932</td>
</tr>
<tr>
<td>YOJ</td>
<td>3.5936</td>
<td>1.0782</td>
<td>0.3642</td>
</tr>
<tr>
<td>JOB</td>
<td>3.3123</td>
<td>1.0022</td>
<td>0.3357</td>
</tr>
<tr>
<td>DELINQ</td>
<td>3.2429</td>
<td>1.9955</td>
<td>0.3286</td>
</tr>
<tr>
<td>MORTDU</td>
<td>2.6823</td>
<td>0.8221</td>
<td>0.2718</td>
</tr>
</tbody>
</table>
Fuel Business Needs with Governance
Visual Lineage of Model and Data

Flexible in innovation and Governed in production
Questions?