Make data work for you with HPE solutions and H2O Driverless AI

HPE ProLiant DL380 Gen10 system and HPE Apollo 6500 Gen10 system help you achieve this
Businesses everywhere have realized that their exclusive data is key to competitive success and now want to put that data to work with artificial intelligence (AI). To scale, data science teams need to adopt new tools and techniques that will allow them to get better results and quickly deliver more insights to the business.

H2O Driverless AI on HPE Apollo GPU-enabled servers is an automatic machine learning platform that gives you an experienced data scientist in a box to create AI-driven products and services to transform your business.

Increasing the business impact of AI by solving a wider variety of business problems is crucial. H2O Driverless AI is optimized to run with the HPE ProLiant and HPE Apollo GPU-enabled servers and automates key portions of the data science workflow. These include feature engineering, parameter tuning, and model optimization to dramatically reduce the time needed to produce accurate models.

**Time series helps forecast sales, predict industrial machine failure, and more temporal predictive use cases**

With the time series capability in Driverless AI, H2O.ai directly addresses some of the most pressing concerns of organizations across industries. This includes use cases such as transactional data in capital markets, track in-store and online sales in retail, and sensor data to improve supply chain or predictive maintenance in manufacturing.

**Trusted AI results**

Delivering machine learning results that a business can trust is an important goal of data science teams. H2O Driverless AI with HPE Apollo GPU enabled servers delivers highly accurate models with machine interpretability that helps explain to the business how the models work. Delivering trusted and transparent results increases adoption of AI and allows your company to comply with government regulations.

**Easy AI deployment**

Model deployment remains one of the most common challenges for data scientists. Models can take weeks or even months to reach production and may be modified to work with production systems. H2O Driverless AI creates ultra-low latency automatic scoring pipelines for easy deployment. In addition, H2O.ai supports training, testing, and model versioning so that data science and business teams can work together to bring models from data science to production in minutes, not months.

And for large-scale enterprise AI deployments, H2O Driverless AI can be combined with BlueData, recently acquired by Hewlett Packard Enterprise. BlueData provides a turnkey container-based software platform to help your company get up and running quickly with distributed AI environments, in a multi-tenant architecture for improved efficiency and reduced costs.
Key features of H2O Driverless AI

AutoViz—exploratory data analysis for Big Data
H2O Driverless AI AutoViz automatically creates data plots based on the most relevant data statistics to help users understand data prior to starting the model building process. This is helpful for data scientists and data engineers who want to better understand the composition of very large data sets and see trends and possible issues such as large numbers of missing values or significant outliers that could impact modeling results.

Automatic feature engineering and model building
Feature engineering is the secret weapon that advanced data scientists use to extract the most accurate results from algorithms. H2O Driverless AI employs a library of algorithms and feature transformations to automatically engineer new, high value features for a given data set. Included in the interface is an easy-to-read variable importance chart that shows the significance of original and newly engineered features. The feature engineering developed for the experiment is explained in the Microsoft® Word Experiment Summary document that is also automatically generated.

Machine learning interpretability (MLI)
H2O Driverless AI provides robust interpretability of machine learning techniques and results. In the machine learning interpretability (MLI) dashboard view, four interactive charts are generated automatically including K-LIME, Feature Importance, Decision Tree, and Partial Dependence Plot. The MLI report also includes techniques such as Shapley values, reason codes, leave-one-covariate-out (LOCO), and MLI scoring pipeline. Each chart and technique help to explore the modeling techniques and results more closely. These techniques are crucial for those who must explain their models to business decision makers, regulators, or customers.

Automatic scoring pipelines
H2O Driverless AI automatically generates both Python and Java scoring pipelines. The latter provides an ultra-low latency automatic scoring pipelines. This is a unique technology that deploys all feature engineering and the winning machine learning model in a highly optimized Python and Java packages that can be deployed anywhere. The technology is critical for enterprises running models that need ultra-fast scoring for real-time applications running on a range of devices.

Identify a use case to get started

<table>
<thead>
<tr>
<th>Financial services</th>
<th>Healthcare</th>
<th>Telecom</th>
<th>Marketing and retail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wholesale and commercial banking</td>
<td>Early cancer detection</td>
<td>Predictive maintenance</td>
<td>Funnel predictions</td>
</tr>
<tr>
<td>• Know your customers (KYC)</td>
<td>• Product recommendations</td>
<td>• Avoidable truck-rolls</td>
<td>• Personalized ads</td>
</tr>
<tr>
<td>• Anti-money laundering (AML)</td>
<td>• Personalized prescription matching</td>
<td>• Customer churn prediction</td>
<td>• Credit scoring</td>
</tr>
<tr>
<td>Card and payment business</td>
<td>• Medical claim fraud detection</td>
<td>• Improved customer viewing experience</td>
<td>• Fraud detection</td>
</tr>
<tr>
<td>• Transaction fraud</td>
<td>• Flu season prediction</td>
<td>• Master data management</td>
<td>• Next best offer</td>
</tr>
<tr>
<td>• Collusion fraud</td>
<td>• Drug discovery</td>
<td>• In-context promotions</td>
<td>• Next best customer</td>
</tr>
<tr>
<td>• Real-time targeting</td>
<td>• ER and hospital management</td>
<td>• Intelligent ad placements</td>
<td>• Smart profiling</td>
</tr>
<tr>
<td>• Credit risk scoring</td>
<td>• Remote patient monitoring</td>
<td>• Personalized program recommendations</td>
<td>• Prediction</td>
</tr>
<tr>
<td>• In-context promotion</td>
<td>• Medical test predictions</td>
<td></td>
<td>• Customer recommendations</td>
</tr>
<tr>
<td>Retail banking</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Deposit fraud</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Customer churn prediction</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Auto-loan</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 2. Use cases by vertical
A complete solution to create trusted machine learning models

H2O Driverless AI is optimized to take advantage of GPU acceleration to speed up the automatic machine learning process and includes support for GPU-accelerated algorithms such as XGBoost, TensorFlow, LightGBM, and more. GPUs allow thousands of iterations for model features and optimizations.

BlueData

H2O.ai has an established partnership and several joint customers with BlueData. The partnership enables HPE customers to realize the full potential of AI and machine learning, helping them to make mission-critical business decisions and deliver data-driven innovation. With the BlueData EPIC software platform, data science teams can instantly spin up Driverless AI running on containers—whether on-premises, in the public cloud, or in a hybrid model.

HPE ProLiant DL380 Gen10 Server

The industry-leading server for multi-workload compute, the secure, resilient 2P/2U HPE ProLiant DL380 Gen10 Server features up to three NVIDIA® Tesla® GPUs and delivers world-class performance and supreme versatility for running AI/ML workloads such as H2O Driverless AI.

HPE Apollo 6500 Gen10 System

This system is a 4U dual-socket server featuring up to eight NVIDIA Tesla GPUs is an ideal high-performance computing (HPC) and deep learning platform. HPE Apollo 6500 Gen10 System provides unprecedented performance with industry-leading NVIDIA GPUs, fast GPU interconnect, high-bandwidth fabric, and a configurable GPU topology to match your workloads.

HPE server technical requirements for running H2O Driverless AI:

**Operating system**
- Red Hat® 7.4/ CentOS 7.4/ SLES 12 SP3/ Canonical Ubuntu 16.04.3

**Libraries:**
- CUDA 9 or 9.2 with NVIDIA drivers greater than 396
- cuDDN greater or equal to 7.2.1 (required only if using TensorFlow)
- OpenCL (required for LightGBM on GPUs)

**Memory per server:**
- 256 GB recommended

**Processor:**
- Dual socket Intel® Xeon® Gold 6254 18 cores 3.1 GHz

**Storage:**
- Recommended that the tmp directory has 500–1 TB of space available. SSD are recommended (preferable NVMe)

**NVIDIA GPU:**
- V100–16 GB or 32 GB recommended

**GPU sizing:**
- Minimum 1 GPU per system
- Recommended: 4–8 GPUs per system
- Typical deployment: 2 GPUs per user

Additional information on sizing requirements can be found here:
- h2o.ai/driverless-ai/latest-stable/docs/userguide/sizing-requirements

Note: H2O.ai Driverless AI and BlueData software are sold separately from the HPE servers mentioned in this document.

Learn more at hpe.com/info/deep-learning