

IBM PowerAI

H2O Driverless AI on Power: AI to do AI

Benoit MAROLLEAU - Cloud Architect
Cognitive Systems
IBM Client Center Montpellier, France
benoit.marolleau@fr.ibm.com



[linkedin.com/in/benoitmarolleau](https://www.linkedin.com/in/benoitmarolleau)



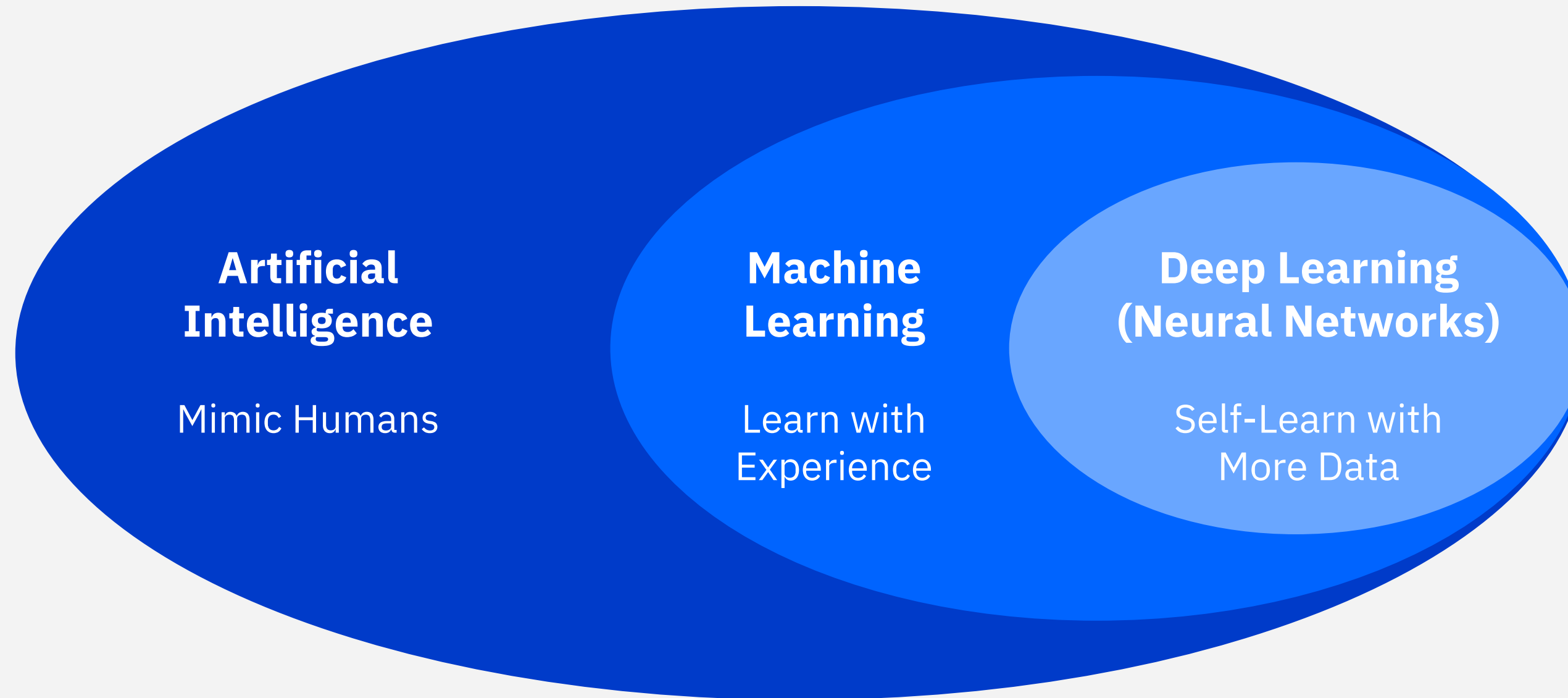
@MarolleauBenoit

Session Objectives

- AI Business
- IBM PowerAI Portfolio
- The IBM Power System AC922 server
- H2O Driverless AI on IBM Power Systems
- H2O Driverless AI on Power vs x86
- Customer Use Cases
- H2O Driverless AI Complements IBM PowerAI Vision

Artificial Intelligence (AI)

IDC 2021 Market Size Projection:
\$14B for AI Servers



57% of AI Market Developed On-Prem

H2O
Driverless AI

PowerAI
Vision

PowerAI

PowerAI
Enterprise

Accelerated
Infrastructure

Auto-ML for non image data
Data Scientists and Non Data Scientists

Auto-ML for Images & Video

Label

Train

Deploy

PowerAI: Open Source ML Frameworks



Large Model Support (LMS)

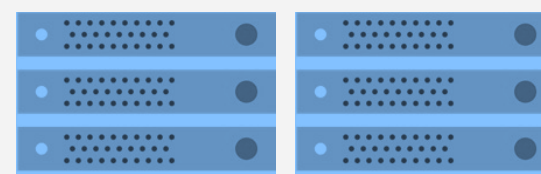
Distributed Deep Learning
(DDL)

Auto ML

IBM Spectrum Conductor with Spark
Cluster Virtualization,
Auto Hyper-Parameter Optimization

**Deep Learning Impact
(DLI) Module**

Data & Model
Management, ETL,
Visualize, Advise



Accelerated Servers



Storage

IBM PowerAI Portfolio

		Deep Learning			ML and DL	Machine Learning
		Power AI Base (WML-CE)	Power AI Enterprise (WML-A)	AI Vision	Watson Studio Local	H2O Driverless AI
Offering	Description	Deep Learning	Deep Learning for the Enterprise	Deep Learning with Video tools	Notebook oriented development environment for ML and DL	Automated Machine learning
	Pricing Model	Free download	Commercial	Commercial	Commercial	Commercial
	Support	Available from IBM	IBM L 1-3 Included	IBM L1-3 Included	Available from IBM	H2O L 1-3
Seller Compensation		Server, Support	Server, Gross SW License	Server, Gross or Term SW License	Server	Server, Net SW License
Applications	Text & Numeric	Yes	Yes	No	Yes	Yes
	Images	Yes	Yes	Yes	Yes	No
	Video	-	Optional add-on	Yes		No
	Primary Persona	Data Scientist	Data Scientist	Line of Business	Data Scientist	Data Scientist
	Second persona	IT	IT	IT	IT	Line of Business
	User Skill Level	High	Medium to high	Low	Medium to high	Low to Medium
	Strengths	Rapid deployment, high performance, scale	enterprise grade, High performance, rapid Deployment	Rapid deployment, simple GUI high performance	Notebook based development environment, strong collaboration, model management	Simplified deployment, intuitive user interface, automatic pipelines, "explainability" for models, end to end automation
Platform	Distributed DL (DDL)	1-4 nodes	1-thousands of nodes	-	-	-
	Large Model Support	Yes	Yes	-	-	-
	Server(s)	S822LC or AC922	S822LC or AC922	S822LC or AC922	S822LC or AC922, LC922	S822LC, AC922, LC921/922
IBM Products	Spectrum MPI (DDL)	Limited to 4 nodes	Included	?		Optional add-on
	Spectrum Conductor DLI	Optional add-on	Included	?	Optional Add On	Optional add-on
	IBM Watson Studio Local	Optional add-on	Optional add-on	No		Optional add-on
Cloud	IBM Cloud Public	Yes	No	Trial only	Watson Studio	?
	IBM Cloud Private	Yes	Yes	Yes	Yes	Yes

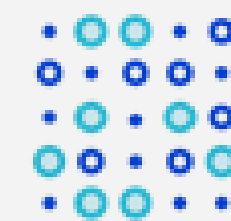
IBM Power System AC922

Accelerated Compute



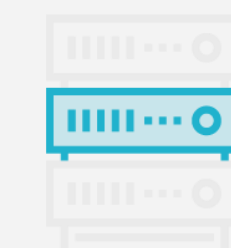
An Acceleration Superhighway

Unleash accelerated computing potential in the post CPU-only era



Designed for the AI Era

Architected for the modern analytics and AI workloads that fuel insights

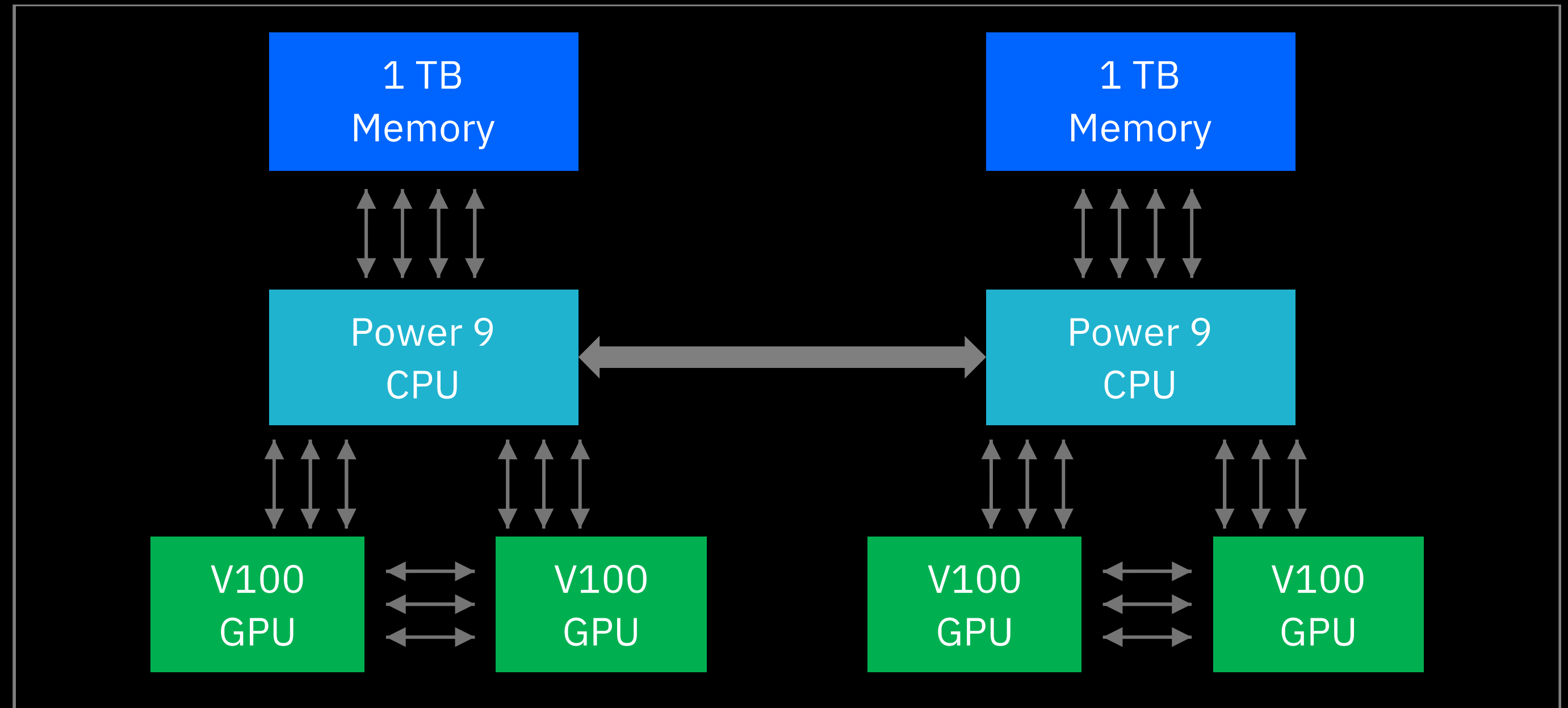
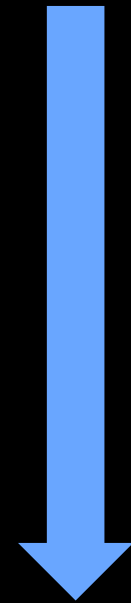


Delivering Enterprise-Class AI

Cutting-edge AI innovation data scientists desire, with dependability IT requires

5x Faster Data Communication with Unique CPU-GPU NVLink High-Speed Connection

Fast Transfer
via NVLink



IBM Power System AC922
Deep Learning Server (4-GPU Config)

H2O Driverless AI on IBM Power Systems

“AI to do AI”

Sri Satish Ambati, CEO and Co-founder, H2O.ai



H2O.ai Company Overview

Company	Founded in Silicon Valley in 2012 Funded: \$75M. Investors: Wells Fargo, NVIDIA, Nexus Ventures, Paxion Ventures
Products	<ul style="list-style-type: none">• H2O Open Source Machine Learning (14,000 organizations)• H2O Driverless AI – Automatic Machine Learning
Leadership	Leader in Gartner MQ Machine Learning and Data Science Platform
Team	100+ AI expertise (7 of the world's top 100 Kaggle Grandmasters/expert data scientists)
Global	Mountain View, London, Prague, India



Worldwide Recognition in the H2O Community

Open source community

222 OF THE 500

H₂O

8 OF TOP 10 BANKS

7 OF TOP 10 INSURANCE COMPANIES

4 OF TOP 10 HEALTHCARE COMPANIES

Commercial Customers

<div>CATALINA</div> <div>DIRECT MAILERS</div> <div>G5</div> <div>acxiom.</div> <div>poder.IO POWER OF PREDICTION</div> <div>Integral Ad Science</div> <div>Nielsen Catalina SOLUTIONS</div> <div>rc</div>	<div>STANLEY</div> <div>BLACK & DECKER</div> <div>H-E-B</div> <div>Travelport</div> <div>Walgreens</div> <div>ebay</div> <div>Booking.com</div> <div>macy's</div>	<div>CREDIT SUISSE</div> <div>WELLS FARGO</div> <div>citi</div> <div>deserve</div> <div>experian.</div> <div>EQUIFAX</div> <div>RBC</div> <div>MarketAxess</div> <div>ING DISCOVER</div> <div>Capital One</div> <div>PayPal</div>	<div>AEGON</div> <div>XCEEDANCE</div> <div>underwrite.ai</div> <div>TRANSAMERICA</div> <div>PROGRESSIVE</div> <div>ZURICH</div>	<div>opta INFORMATION INTELLIGENCE</div> <div>KAISER PERMANENTE</div> <div>CHANGE HEALTHCARE</div> <div>ARMADA Health Care</div> <div>aetna</div>	<div>starling</div> <div>ADP</div> <div>pwc</div>	
HW Vendors	Marketing	Retail	Financial	Insurance	Healthcare	Advisory & Accounting

Driverless AI Demand

3300+

Trial Requests

3M+

Models Built

“H2O.ai's reference customers gave it the highest overall score for sales relationship and overall service and support” - Gartner MQ 2018

MC06 / November 14, 2018 / © 2018 IBM Corporation

10

H2O.ai is a Leader in the 2018 Gartner Data Science and Machine Learning Platforms Magic Quadrant

- Technology leader with most completeness of vision
- Recognized for the mindshare, partner network and status as a quasi-industry standard for machine learning and AI
- H2O.ai customers gave the highest overall score among all the vendors for sales relationship and account management, customer support (onboarding, troubleshooting, etc.) and overall service and support

Figure 1. Magic Quadrant for Data Science and Machine-Learning Platforms



Source: Gartner (February 2018)

The Challenges of Enterprise AI Adoption

Data is a Team Sport

~100

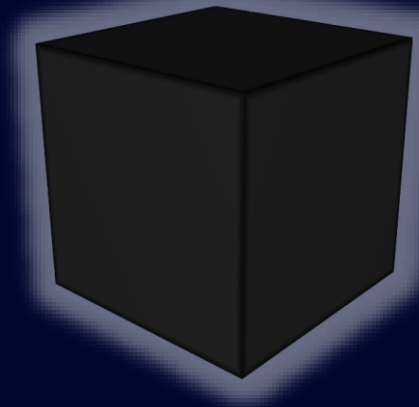
Data science experts in the world

Lack of AI Talent

Months

Time for a data scientist to build a model

Time to Insights slow



Black box models

Lack of Trust in AI

"US alone faces a shortage of 190,000 people with analytical expertise."

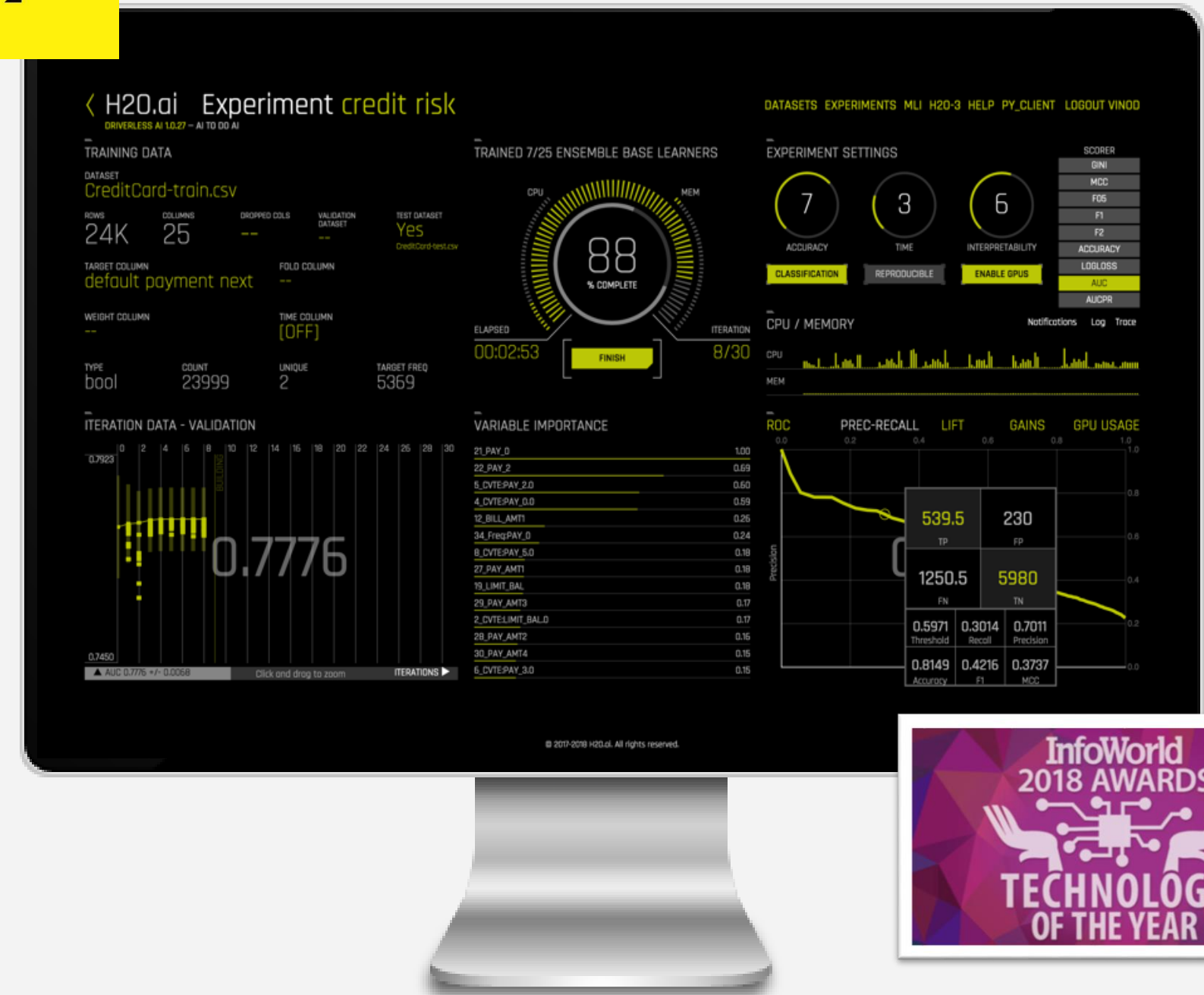
McKinsey&Company

H₂O.ai

H2O Driverless AI Delivers “Expert Data Scientist in a Box”

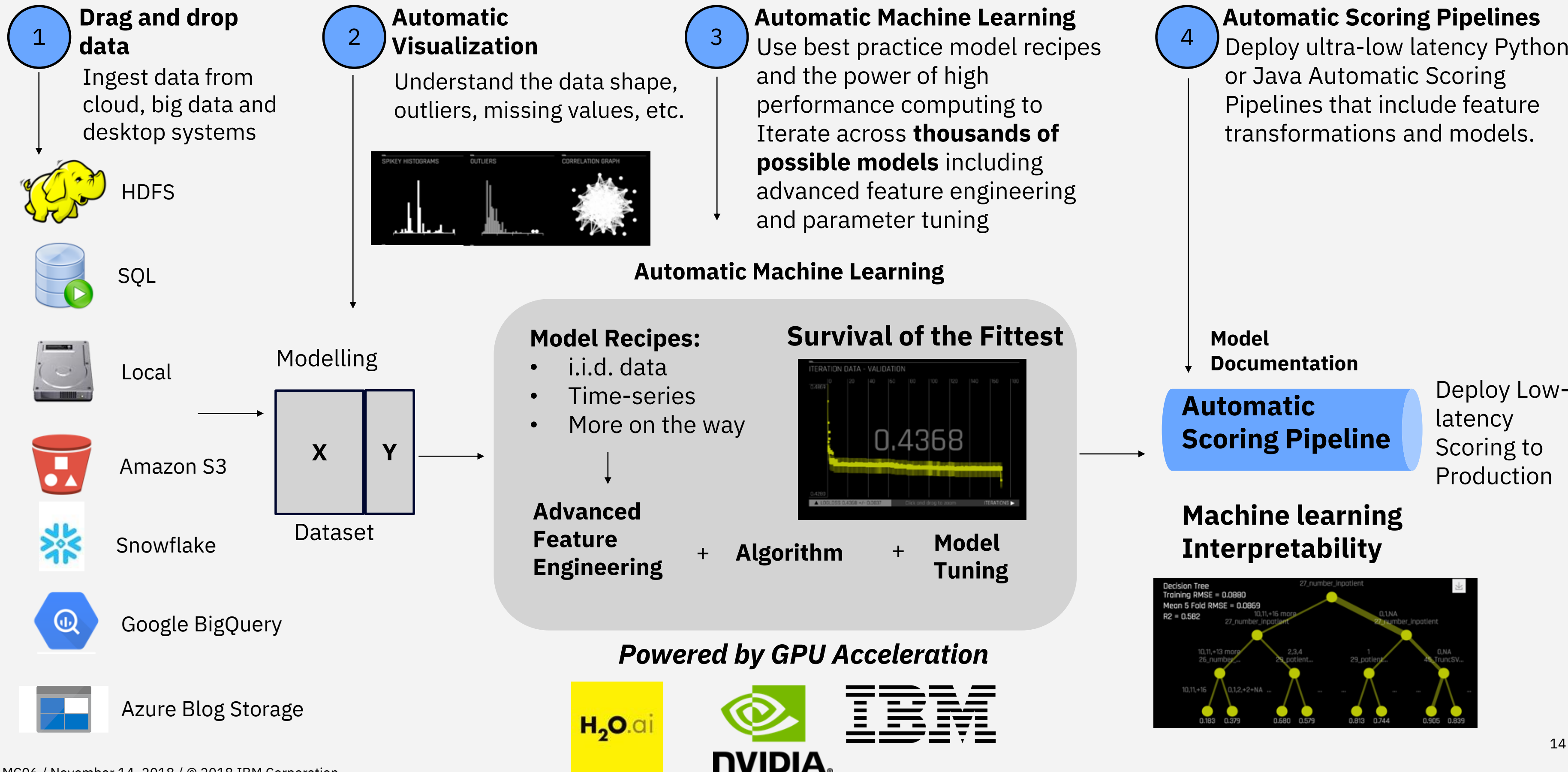
- Award-winning software
- Created and supported by world renowned AI experts from H2O.ai
- Empowers companies to accomplish AI and ML with a single platform
- Performs the function of an expert data scientist and adds more power to both novice and expert teams
- Details and highlights insights and interpretability with easy to understand results and visualizations

H2O.ai



21 day free trial for [Driverless AI](#)

H2O Driverless AI: How it Works



Driverless AI Delivers “Expert Data Scientist in a Box”

Automation:

- Visualization
- Feature Engineering
- Model Tuning
- Time Series

Deployment:

- Automatic Pipelines
- Low latency inferencing

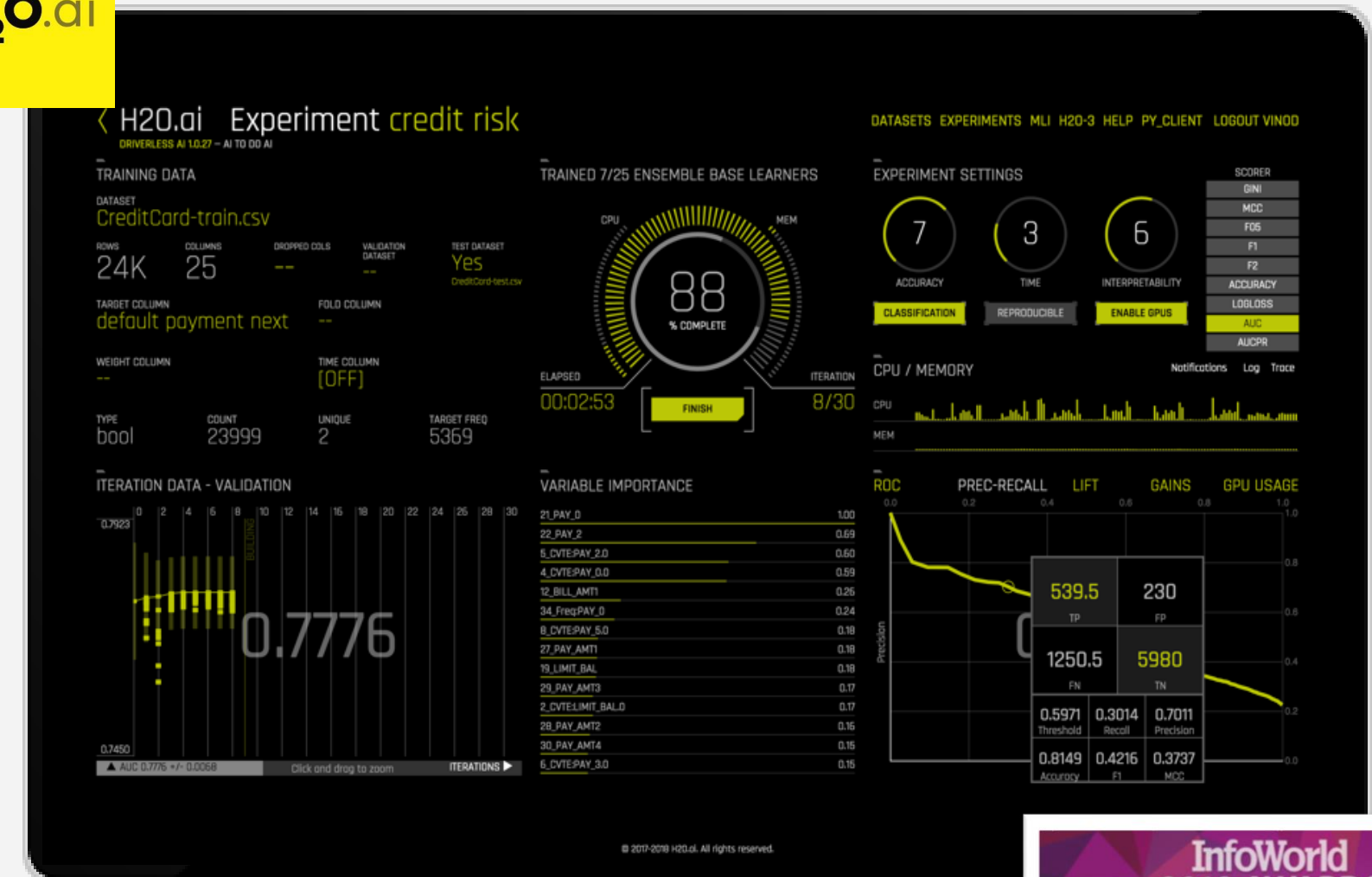
Trust:

- Explainability and reason codes
- Automatic documentation

Enterprise Ready:

- Secure – LDAP, Kerberos
- Scalable – Scale with GPUs
- Support for enterprise data sources

H₂O.ai



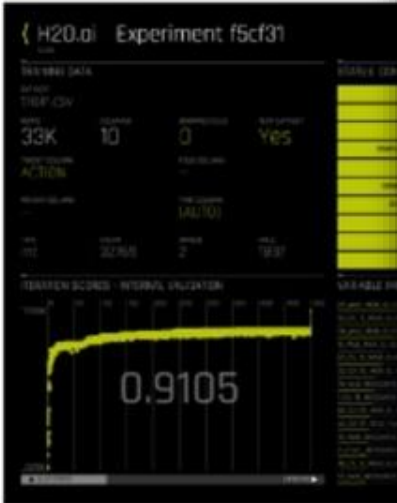
21 day free trial for [Driverless AI](#)

Highly Accurate, Easy to Use, Really Fast

- Simple interface
- Automatic feature engineering to increase accuracy
- Automatic recipes for solving wide variety of use-cases
- Automatic tuning to find and tune the right ensemble of models

Driverless AI: top 5% in Amazon Kaggle competition

Driverless AI produces feature engineering pipeline (“more columns”) for downstream use



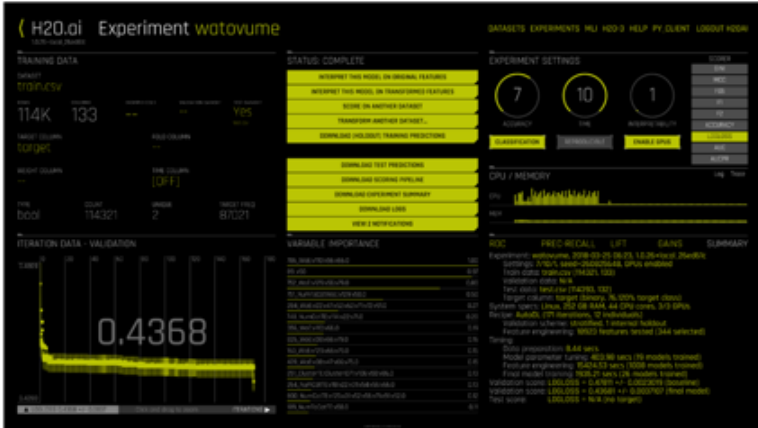
Amazon.com - Employee Attrition Prediction

Predict an employee's access to the company's internal network

\$5,000 · 1,687 teams · 4 years ago

Driverless AI: 80th place (out of 1687 - top 5%)

Driverless AI: Top-10 in BNP Paribas Kaggle competition



single run, fully automated: 2h on DGX Station! 6h on PC

BNP Paribas Cardif Claims Management

Can you accelerate BNP Paribas Cardif's claims management process?


\$30,000 · 2,926 teams · 2 years ago

Submission and Description	Private Score	Public Score
sub.csv 2 minutes ago by Arno Candel 940b9f 7/10/1 cv 0.4354 finished after 172 iters	0.42945	0.43156

Driverless AI: 10th place in private LB at Kaggle (out of 2926)

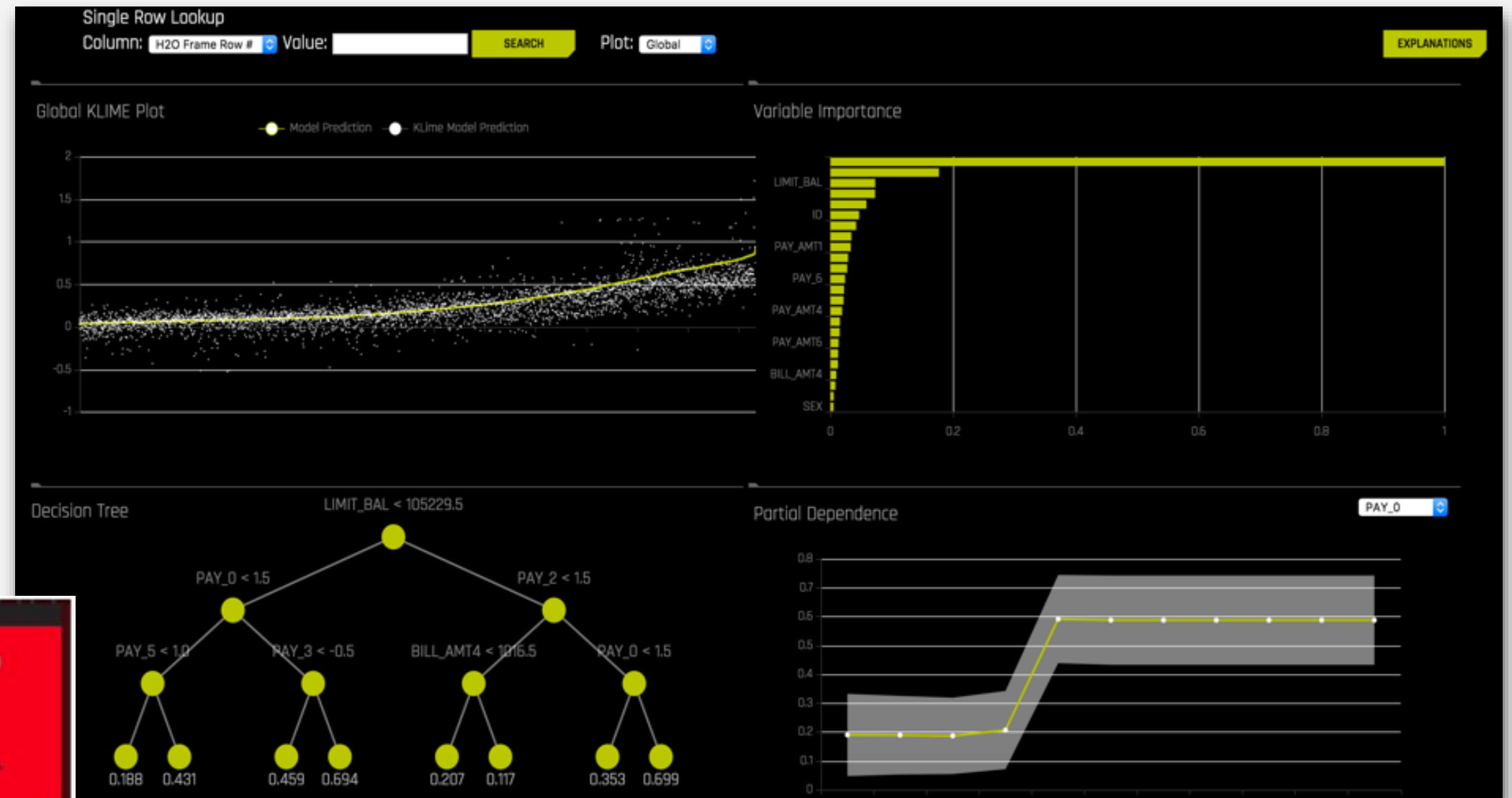
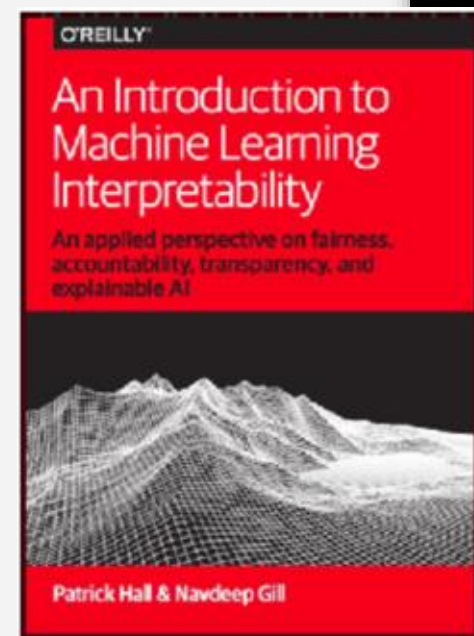
2 months for Grandmasters — 2 hours for Driverless AI

#	Δpub	Team Name	Kernel	Team Members	Score @	Entries	Last
1	—	Dexter's Lab			0.42037	198	2y
2	—	escalated chi			0.42079	162	2y
3	—	Exploding Kittens			0.42182	124	2y
4	—	Branden Nickel utility			0.42259	251	2y
5	—	the flying burrito brothers			0.42450	264	2y
6	—	n_m			0.42535	4	2y
7	—	PAFY			0.42557	310	2y
8	—	KAME			0.42688	121	2y
9	—	Jack (Japan)			0.42744	22	2y
10	▲1	Dmitry & Bohdan			0.43000	192	2y
11	▲1	Li-Der			0.43006	56	2y
12	▼2	BK3M2PRS			0.43089	338	2y
13	—	x2x4x8			0.43107	55	2y
14	—	Frenchies			0.43146	134	2y
15	▲1	Ains			0.43168	55	2y
16	▼1	maze runners			0.43262	164	2y
17	—	BLR-2			0.43313	129	2y
18	▲3	no one			0.43317	88	2y



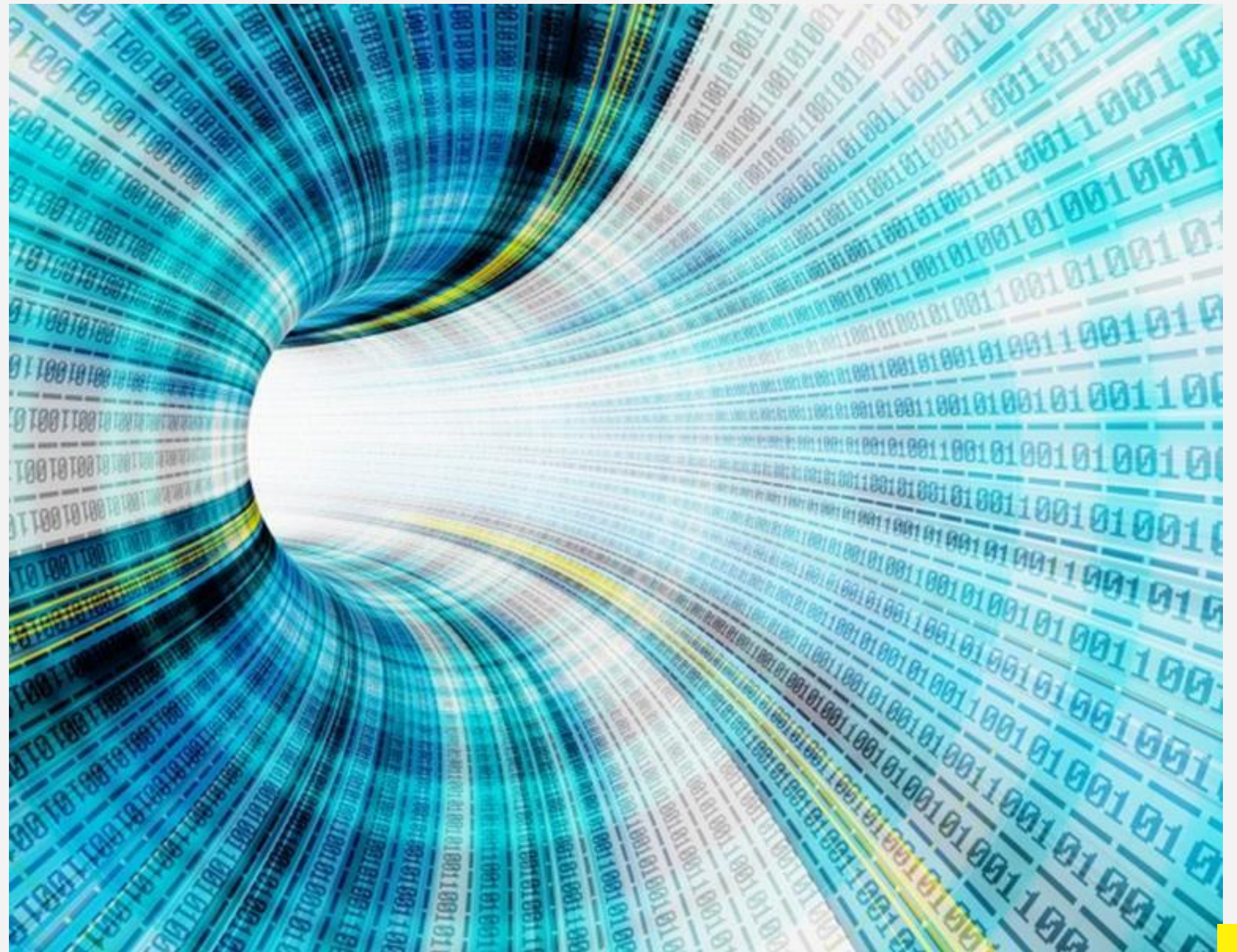
Industry Leading Interpretability for Trust and Compliance

- Interpretability for debugging, not just for regulators
- Get reason codes and model interpretability in plain English
- K-Lime, LOCO, partial dependence and more



Easy Deployment for Low-latency Models

- Stand-alone scoring pipeline that is easy for IT to deploy and manage including all feature transformations
- Easy to update when a new model version is available
- Optimized the scoring code for complex (big data) models
- Streamlined scoring code to deploy on any device on the edge, mobile, etc.
- Really fast (millisecond) to satisfy today's real-time apps



Proof points for Power vs x86

H2O Driverless AI on IBM Power Systems - AI in a Box

2.6x

More RAM

Big Data Scale

9.5x

Max I/O bandwidth

High Speed Data Transfer

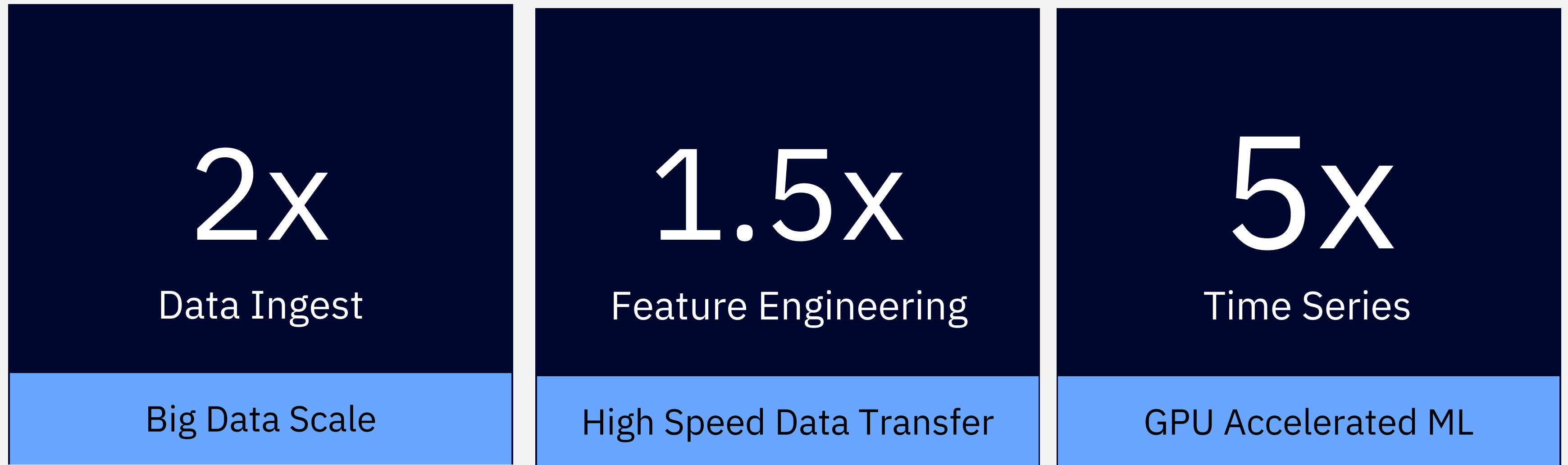
30x

Faster on GPUs

GPU Accelerated ML

Integrated Systems Approach

H2O Driverless AI on IBM Power Systems - AI in a Box



Integrated Systems Approach

Customer Use Cases

The logo for H2O.ai is displayed within a solid yellow square. The text "H2O.ai" is rendered in a bold, black, sans-serif font. The "H" and "2" are black, while the "O" is a lighter shade of gray, and ".ai" is also in a lighter shade of gray.

H₂O.ai

H2O.ai at Capital One: 25 Use Cases

Card/Payments Business

- **Application Fraud Prediction**
Reduction of Fraudulent Applications
- **Transaction Fraud Prediction**
Reduce loss and Improving Customer Experience
- **Real-time Targeting**
Improve Conversions
- **Credit Risk Scoring**
Optimizing credit portfolio
- **In-Context Promotion**
Improving Customer Experience and Conversions

Consumer Banking

- **Deposit Fraud Prediction**
Faster customer access to funds, Improved Customer Experience
- **Customer Churn Prediction**
Improved Customer Retention
- **Auto-Loan Credit Risk**
Optimizing credit portfolio
- **Proof of Income Prediction**
Reducing manual labor and verification
- **Direct-Mail Marketing**
Marketing Optimization

Wholesale / Commercial Banking

- **Know Your Customers (KYC)**
Probabilistic Matching Engine for Customer 360 view
- **Anti-Money Laundering (AML)**
Reduce False Positives, Suspicious Activity Report (SAR's) filing
- **Document Classification**
Improved Machine Learning based document classifier, reducing of manual labour.



Infrastructure

- **Security Cyberlake**
Windows Syslog / Domain Generation Algorithm
- **Autonomic Computing**
Improving operational efficiency of business critical systems.
- **Master Data Management**
Single Master Customer record with Probabilistic Matching Engine
- **Model Management & Governance**
Model Management, Documentation
- **Tech College**
Training and Evangelism



Financial Fraud Detection



- Driverless AI matched **10 years** of expert feature engineering
- Increased accuracy from **0.89 to 0.947 (6%)** in detecting fraudulent activity
- **6X** speed up when using H2O4GPU with Driverless AI

Experiment

- Training time (subset of data) – Driverless AI on GPU 6x faster
 - laptop (accuracy 1) - ~ 2 hours
 - GPU (accuracy 1) – 21 minutes; (accuracy 5) – 58 minutes

ID	TARGET	TRAIN SCORE	TEST SCORE	SCORER	ACCURACY	TIME	INTERPRETABILITY	STATUS	TIME
13c6co	is_cc_bad	0.94703	NA	AUC	1	1	1	Done	01:53:29
067d32	is_cc_bad	0.94773	NA	AUC	5	5	5	Done	00:58:39
e55a93	is_cc_bad	0.94658	NA	AUC	1	1	1	Done	00:21:59

PayPal © 2017 PayPal Inc. Confidential and proprietary.

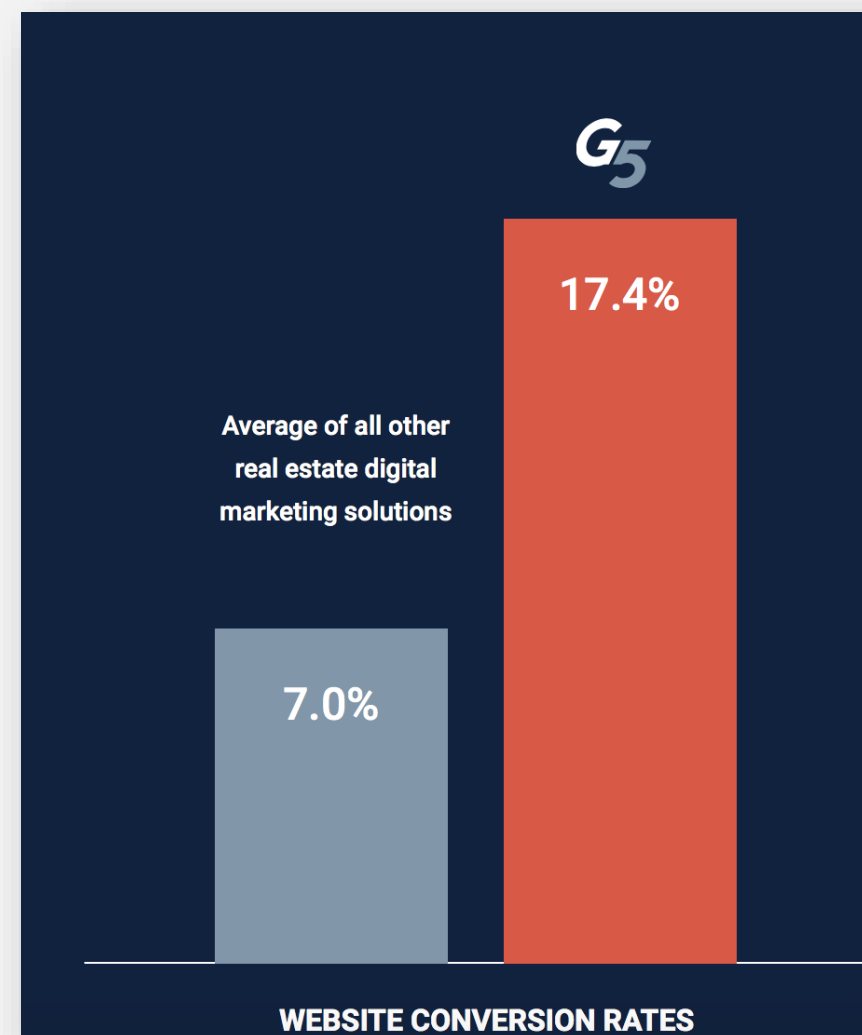
“Driverless AI is giving amazing results in terms of feature and model performance “

Venkatesh Ramanathan
Senior Data Scientist, PayPal

Marketing Optimization for the Real Estate Market



- Outperforms other real estate digital marketing solutions by **2.5X**
- A G5 client saved **\$500K** annual digital spend while increasing web traffic **3X**
- **10X** faster model creation



“Driverless AI helped us gain an edge with our **Intelligent Marketing Cloud** for our clients. AI to do AI, truly is improving our system on a daily basis.”

Martin Stein
Chief Product Officer

Improve Manufacturing Sales and Forecasting

StanleyBlack&Decker

- Time savings of **25%** with 1 data scientist
- Saved **one month** of time in model tuning and training for industrial product line
- Accurately predicted supplies and materials for a future client order increasing forecast accuracy



“H2O Driverless AI feature engineering is better than anything I've seen out there right now. And the scoring pipeline generation is probably one of the bigger pluses for me. It's a massive time saver.”

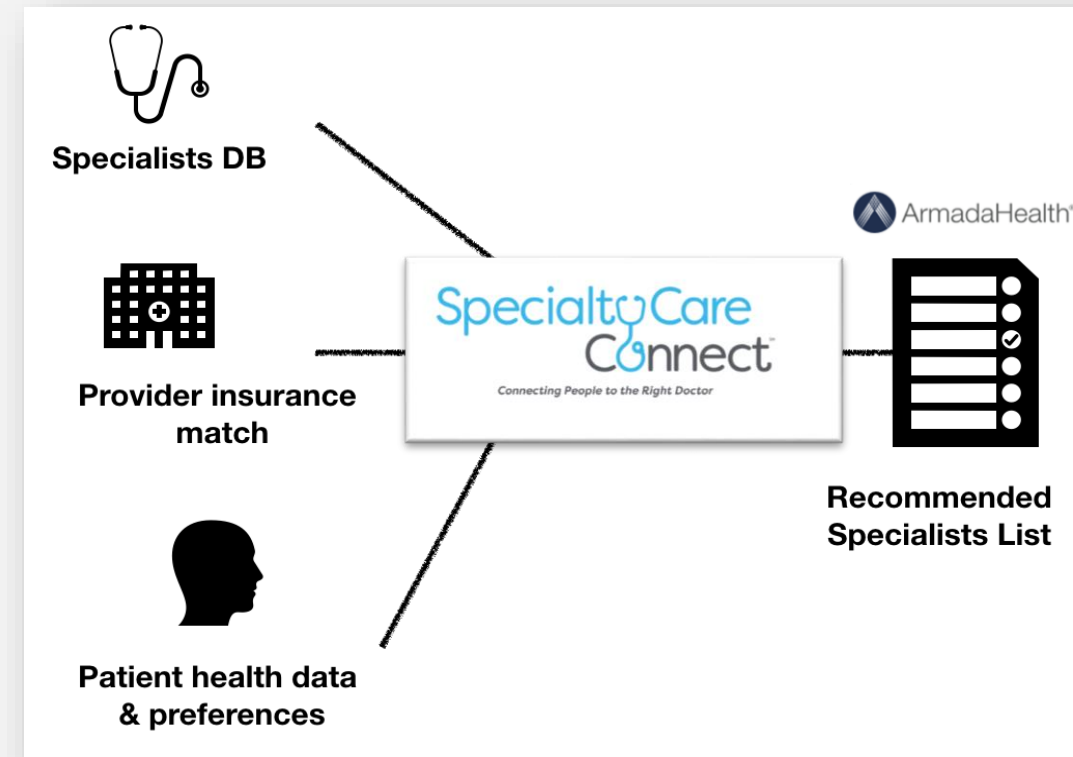
Robert Coop
Sr. Data Scientist
Stanley Black & Decker

H₂O.ai

Connecting Patients to Specialists for Better Healthcare



- Companies have seen “skyrocketing” net promoter scores and “near perfect” customer satisfaction rates
- Customer loyalty and premium retention rates have increased
- Reduces costs, while patients receive care faster



“Driverless AI powers our data science team to operate efficiently and experiment at scale. With this latest innovation, we have the opportunity to impact care at large.”

Bharath Sudarshan
Director of Data Science and Innovation
Armada Health

H2O Driverless AI Delivers Value in Every Industry

H₂O.ai



Financial Services

**+6%
Accuracy**

Matched 10 years of Machine Learning Expertise

“Driverless AI is giving amazing results in terms of feature and model performance “

*Venkatesh Ramanathan
Senior Data Scientist,
PayPal*



Healthcare

**Near perfect
scores**

Increased customer satisfaction

“Driverless AI powers our data science team to operate at scale. We have the opportunity to impact care at large.”

*Bharath Sudarshan
Director of Data Science,
Armada Health*



Marketing

**2.5X
Performance**

Outperforms alternative digital marketing

“Driverless AI helped us gain an edge for our clients. AI to do AI, truly is improving our system on a daily basis.”

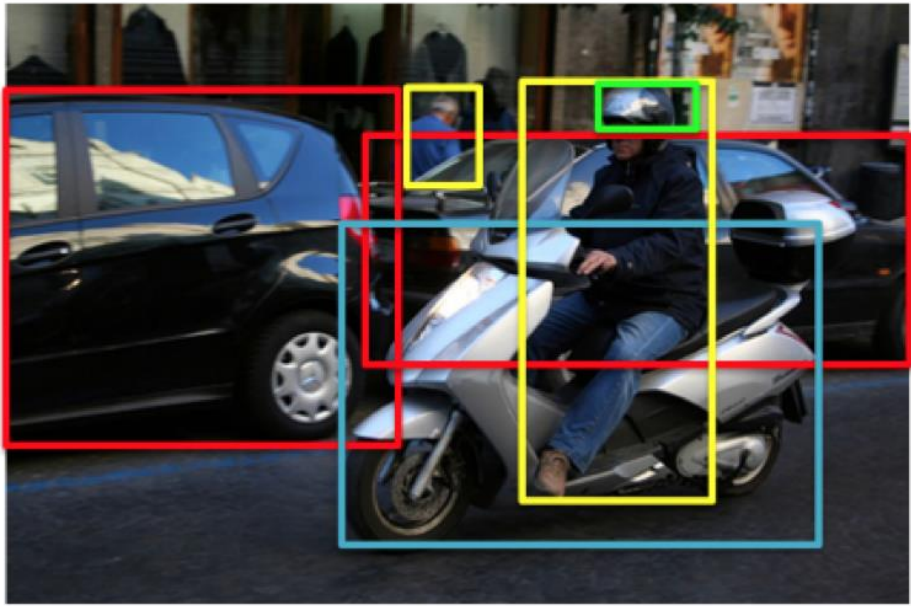
*Martin Stein
Chief Product Officer,
G5*

H2O Driverless AI & IBM PowerAI Vision

H2O Driverless AI Complements IBM PowerAI Vision



IBM Power AI delivers Deep Learning for Images



Person
Car
Motorcycle
Helmet

Facial Insights

Press "Take Photo" to select a photo:



Your gender: Male



MC06 / November 14, 2018 / © 2018 IBM Corporation



H2O Driverless AI is an Automatic Machine Learning

Transactional Data: Store Level

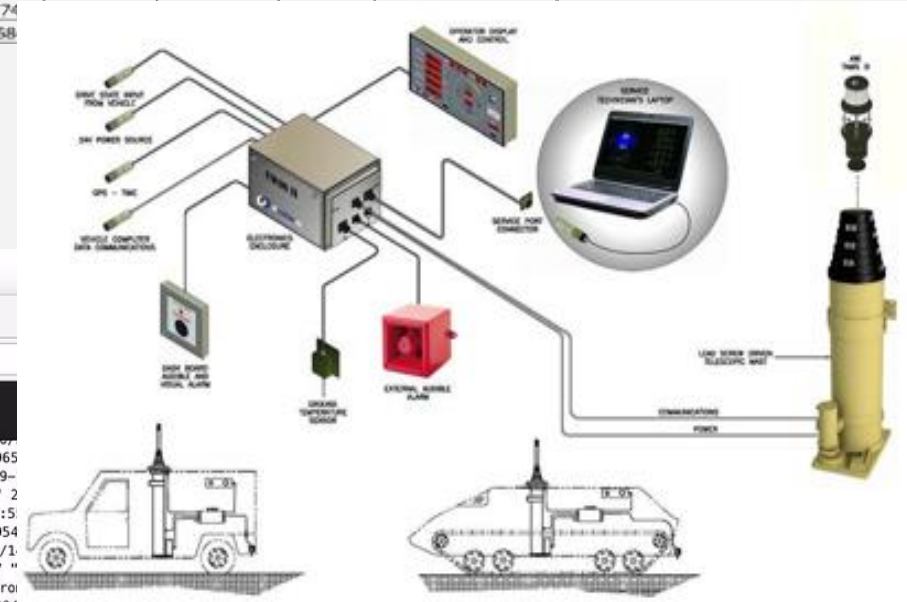
Transactional

Product Code	Amount
G324516283	2.3
5820257187	1.4
P0322773511	1.0
T835138161	2.1
X363594534	1.2
A293943851	3.8
B217229545	1.4
TOTAL	14.4

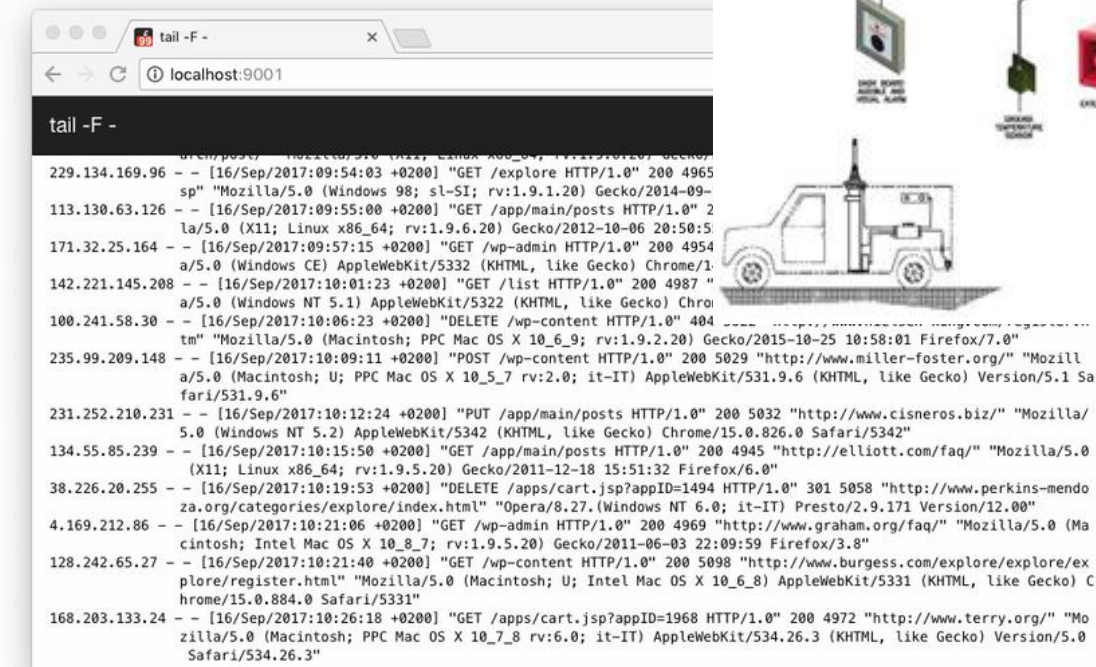
Example: Flat File

Transaction ID	Date	Time	AMOUNT	Card_number
p3913045367	11/24/2010	3:20:32 AM	14.47525	*****2323
q6781991970	12/4/2010	10:30:10 AM	39.46618	*****6451
r9178287202	11/2/2010	1:23:21 AM	24.99964	*****2179
s7329082178	10/8/2010	12:29:40 AM	19.21324	*****5826
t6835491952	10/17/2010	10:36:12 PM	39.56086	*****9498
u8333871154	12/5/2010	5:22:59 PM	33.35401	*****9379
v1240343519	10/27/2010	4:12:38 AM	49.64481	*****5466
w5406123613	11/26/2010	10:36:36 PM	24.23247	*****1816
x8906115216	11/3/2010	1:57:33 PM	32.45101	*****2662
y1400944560	11/25/2010	7:26:45 PM	32.12293	*****4330
z3665080889	10/9/2010	11:09:58 AM	21.93351	*****6533
p6738256686	11/23/2010	10:14:36 AM	21.71996	*****4615
q8964299443	11/24/2010	1:19:24 AM	15.46741	*****7694
r9012945025	11/22/2010	2:02:15 PM	31.14201	*****7140
s2116305133	11/23/2010	7:15:13 PM	43.16047	*****9208
t478724264	12/2/2010	12:08:46 PM	40.14018	*****2695
u9844400319	10/11/2010	3:10:19 AM	23.70907	*****2980
v2388798974				
w567236458				

Sensors



Log



Call of Action

You are invited to join us tomorrow to our
Dallas IBM PowerAI Meetup
Thursday Nov 15, 18:00-20:00 at XXXXX
Sponsored by IBM and H2O.ai

For a free 21-day trial version, visit the following link:

<https://www.h2o.ai/try-driverless-ai/>

For information about the partnership of IBM and H2O.ai, visit the following link:

www.h2o.ai/ibm

