Power Week

Université IBM i 2019

22 et 23 mai

IBM Client Center Paris

i for Business

S12 - PowerAI, Watson et IBM i – Partie 2 Demonstrations

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

linkedin.com/in/benoitmarolleau
@MarolleauBenoit

Online version – EN & FR
https://ibm.biz/bma-wiki

Agenda

- PARTIE 1
  - Presentation bientôt disponible sur https://www.2019.universite-i.fr/ et ici https://ibm.biz/bma-wiki

- PARTIE 2

  Exemple et demonstration – Technologies en action sur IBM i

  Machine Learning sur IBM i: Scikit-Learn

  Machine Learning accéléré sur H2O.ai Driverless AI

  Deep Learning (Visual Recognition) avec PowerAI Vision, IBM i et Node.js

IBM i, Machine Learning & Solutions IBM

Quelques questions à se poser

Machine Learning use cases & relevance for business ?

- Business Analyst & Data Expertise
- Model Precision / Accuracy
- Model evaluation, monitoring & re-training

Deployment Options ? (vs. regulation, cost, performance, skills, other technical aspects)

- AI in the Public Cloud
- AI in the Datacenter (on premise)
- AI on the edge

Data Science Phases ?

- Data Preparation & Model building (which data sources, language & EDI, framework, libraries, algorithm?)
- Model training & validation
- Model evaluation, deployment for Inference (REST API? Batch? Streaming?)

Performance / Cost / Time to market ?

- Accelerated ML (GPU, FPGA,...) or not for model training? Model Inference?

Regulation & Auditability vs. model understanding ?

- Model Fairness
- Model Interpretability

IBM i, Machine Learning & Solutions IBM

Solutions Cloud ou dans le datacenter

1

ML Capabilities on IBM i
Simple, Efficient
No GPU, runs on CPUs

Machine Learning Libraries

2

PaaS/SaaS model, Pay per use
Great functionalities, quickly available, CPU and GPU

IBM Watson

Watson Studio + WML + Data Platform

IBM Cloud

Advanced users (i.e. Cluster with 4+ PowerAI)
- WML-A for Hyperparameter optimization
- Unique Linear Scalability : Distributed DL (DDL)
- State of the art EGO job Scheduler from HPC

3

python

PowerAI Vision

H_2O.ai

Watson Studio +WML

Plugin – Training & Inference

WML-A

IBM Watson

AC922

PowerAI Base (WML-CE) – optional ICP K8s Deployment

docker kubernetes

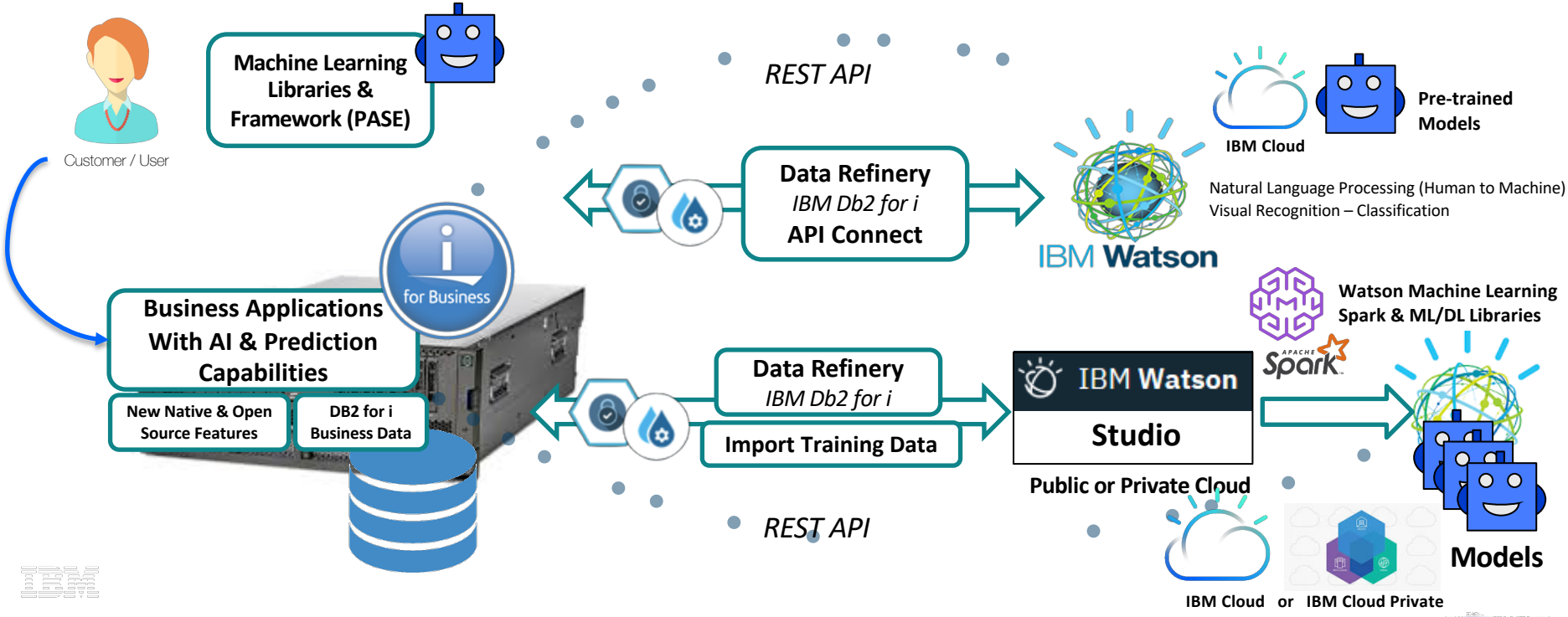
AC922 AC922

Accelerated ++
Divide Training time by x vs. x86
DL: Large Model Support & NVLink
ML: Accelerated ML with SnapML and Nvidia Rapids
Free PowerAI Supported & Optimized Frameworks –
Simple Docker based deployment. Optional K8s/ICP

On top of PowerAI Base (WML-CE):
Watson Studio, H2O Driverless AI, PowerAI Vision…
Get Results in minutes

IBM i & Artificial Intelligence

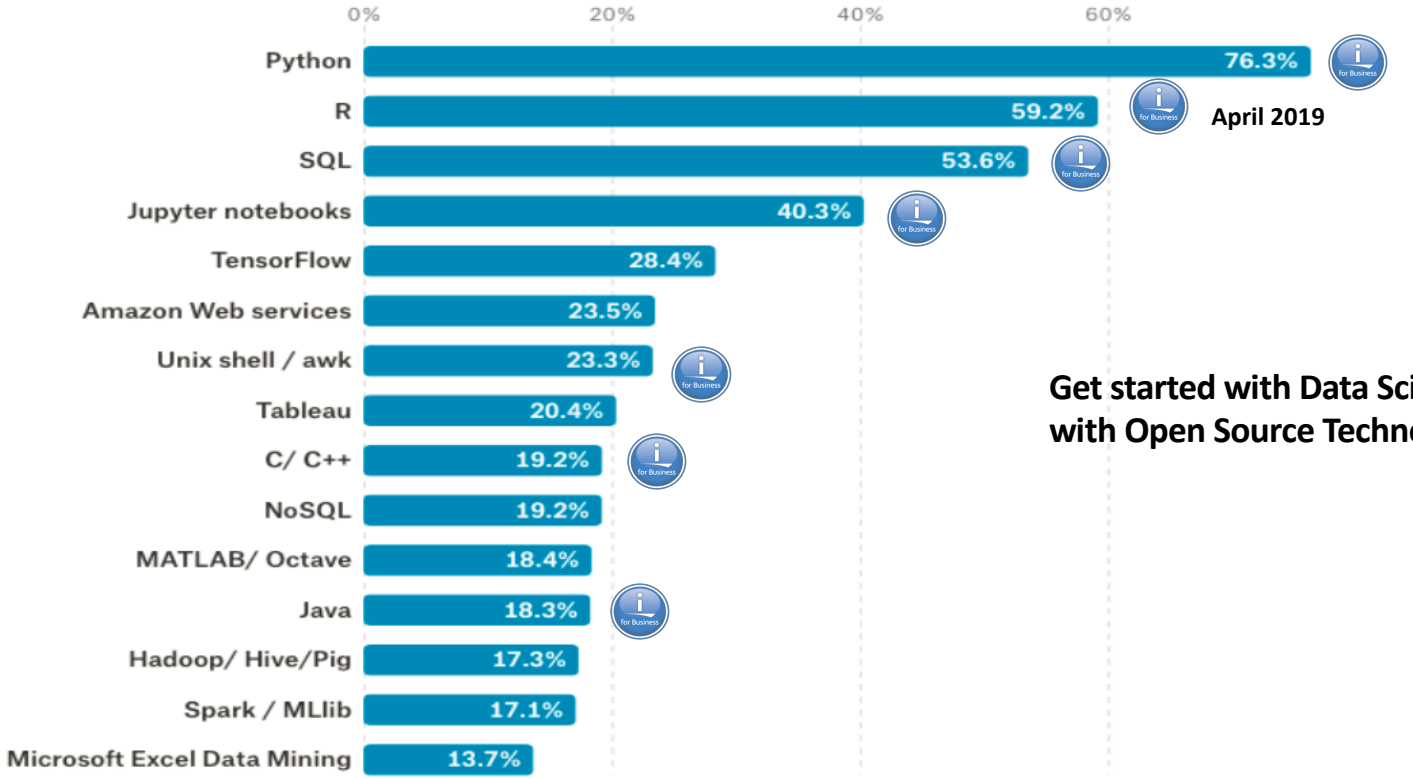
Scenario 2: Utiliser Watson Developer Cloud (API) , IBM Watson Studio (Cloud ou Local)



- Data is the key in all AI projects: your business data resides on IBM i
- Use pre-trained & customizable models with IBM Watson (Developer Cloud) services in IBM Cloud
- Build your own use case & business specifics models with IBM Watson Studio - IBM Cloud / on premises (Studio Local w/ Cloud Private)

Data Science tools & technologies

Kaggle 2017 Data Science Tools Survey



April 2019

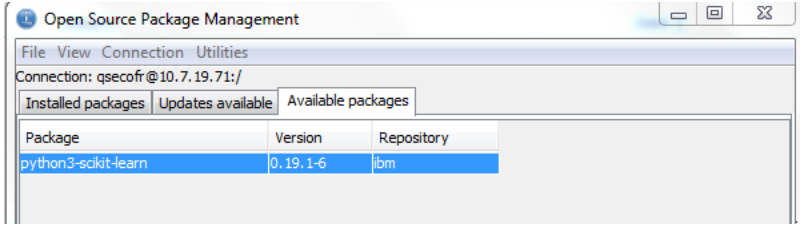
Get started with Data Science on IBM i with Open Source Technologies

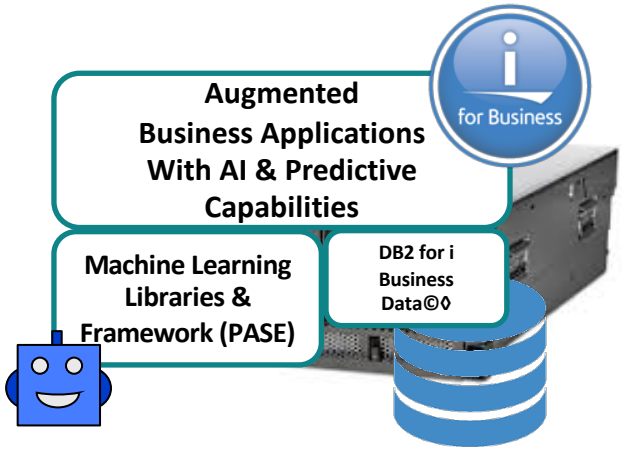
IBM Data Science Technologies & IBM i

Scenario 1: Utiliser les frameworks et langages disponible sur IBM i 7.2+

Data & Scientific Packages Available

- Numpy, Pandas : Data Processing
- **Scipy, Scikit-Learn**
- IPython : interactive Python
- NLTK : Natural Language Processing
- Matplotlib, jupyter : Data Visualization
- R Language (Interpreter, Runtime)
- More to come? ☺





Alternative vs. dataset size, investment on ML etc.

- Training & Inference on IBM i
- Training on PowerAI / WML , Inference on IBM i (<->) (Public or Private Cloud)
- Data preparation on IBM i, Training & Inference on Accelerated Servers (PowerAI, WML-A)

IBM Data Science Technologies & IBM i

Demo1

- Customer Churn Demo
  - Supervised model – classification with Scikit-Learn

- Install yum packages and git clone
  - https://github.com/bmarolleau/firstdemo-scikitlearn-ibmi/

IBM Data Science Technologies & IBM i

Demo1

Classification

Identifying to which category an object belongs to.

Applications: Spam detection, Image recognition.
Algorithms: SVM, nearest neighbors, random forest, …

— Examples

Regression

Predicting a continuous-valued attribute associated with an object.

Applications: Drug response, Stock prices.
Algorithms: SVR, ridge regression, Lasso, …

— Examples

Clustering

Automatic grouping of similar objects into sets.

Applications: Customer segmentation, Grouping experiment outcomes
Algorithms: k-Means, spectral clustering, mean-shift, …

— Examples

Dimensionality reduction

Reducing the number of random variables to consider.

Applications: Visualization, Increased efficiency
Algorithms: PCA, feature selection, non-negative matrix factorization.

— Examples

Model selection

Comparing, validating and choosing parameters and models.

Goal: Improved accuracy via parameter tuning
Modules: grid search, cross validation, metrics.

— Examples

Preprocessing

Feature extraction and normalization.

Application: Transforming input data such as text for use with machine learning algorithms.
Modules: preprocessing, feature extraction.

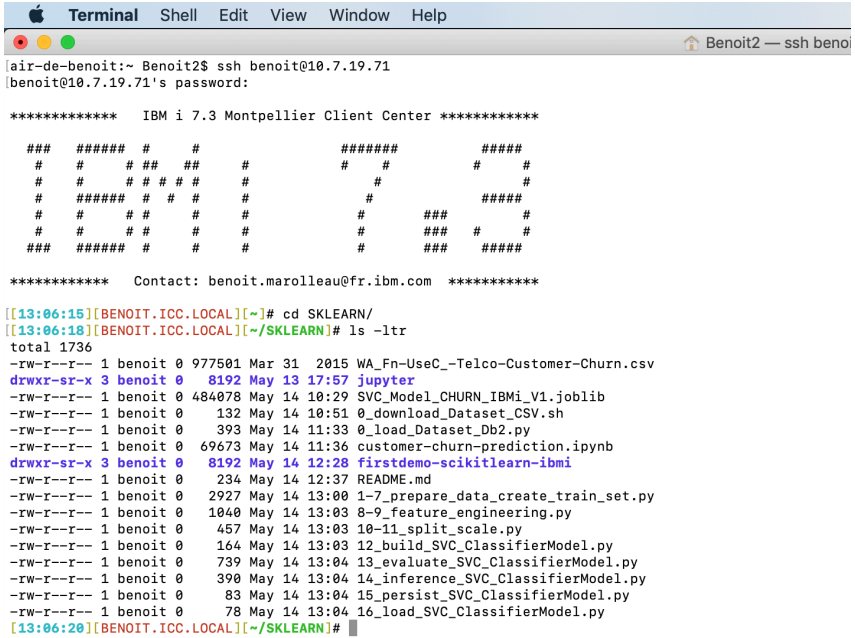
— Examples

IBM Data Science Technologies & IBM i

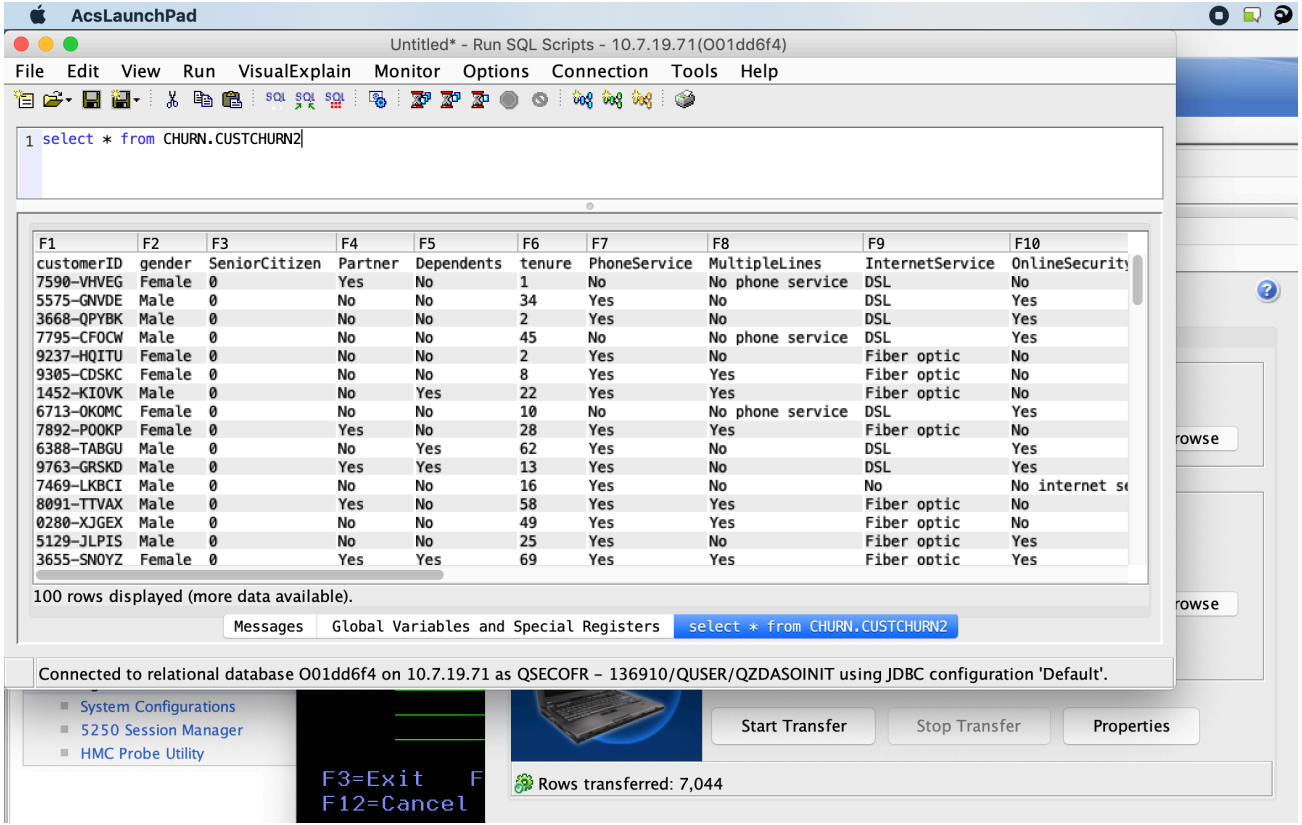
Demo1

- Customer Churn Demo
  - Customer segmentation :
    is this customer going to leave now? soon? or not?
- Dataset from Db2 CHURN.CHURNCUST2
- Use ACS for data transfers (CSV <-> Db2)
- Ipython and/or Jupyter for coding
- From Data preparation to model inference on IBM i

- Final model can infer on IBM i or can be externalized on an accelerated node or Watson ML (On prems, Cloud)



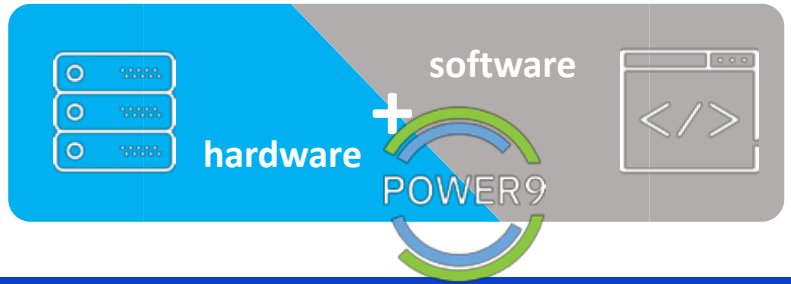
IBM Data Science Technologies & IBM i



Demo 2 & 3 : PowerAI Vision / Accelerated AutoML avec Driverless AI

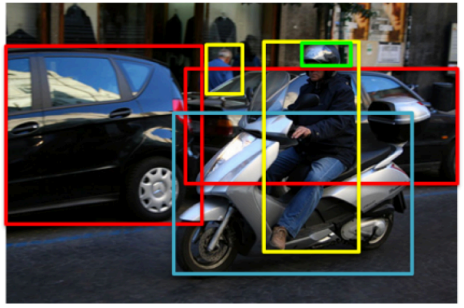
Scenario 3: Utiliser l'état de l'art du Private Accelerated ML/DL avec PowerAI / WML-A,
s'intégrant automatiquement avec des solutions comme Watson Studio Local, PowerAI Vision, Driverless AI.

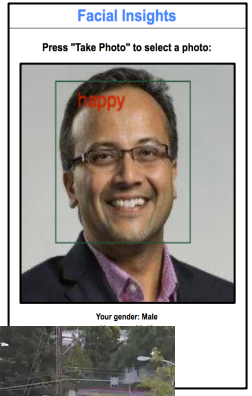
- L'objectif est de ne plus payer à l'usage, d'être autonome, et d'avoir une techno pérenne et performante
  - Prix compétitifs (vs. TCA/TCO, vs. offres Cloud ou « on premise » concurrentes)
- Solution Matérielle et logicielle Open source, supportée par IBM.
  - Solution designée pour le Machine Learning et Deep Learning accéléré. Base du supercalculateur CORAL Summit
  - Disponible « on premise » , dans le Cloud public.
- Facilement Cloudifiable en Cloud Privé via Kubernetes / IBM Cloud Private
  (Hébergeurs, Consolidation des environnements AI)



Demo 2 & 3 : PowerAI Vision / Accelerated AutoML w/ Driverless AI

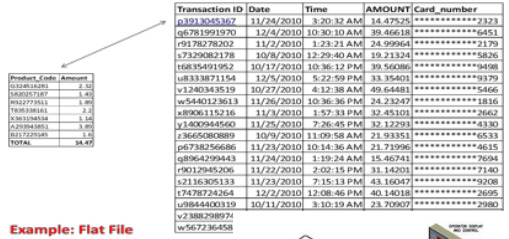
IBM Power AI delivers Deep Learning for Images



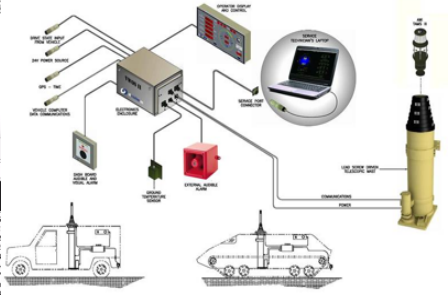


H2O Driverless AI is an Automatic Machine Learning

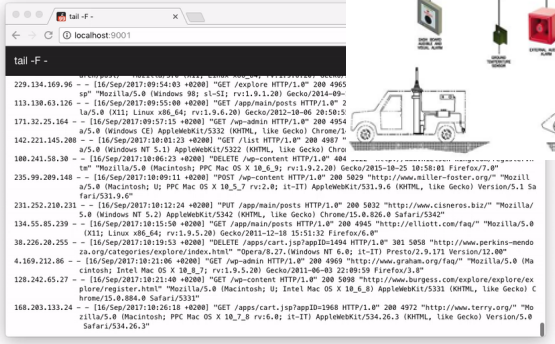
Transactional



Sensors



Log

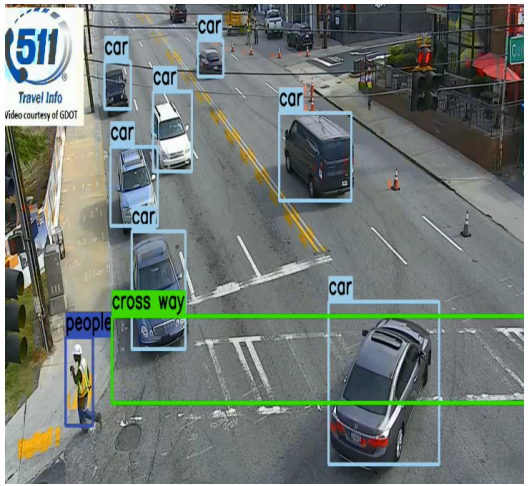


Comparing AI Offerings on Power

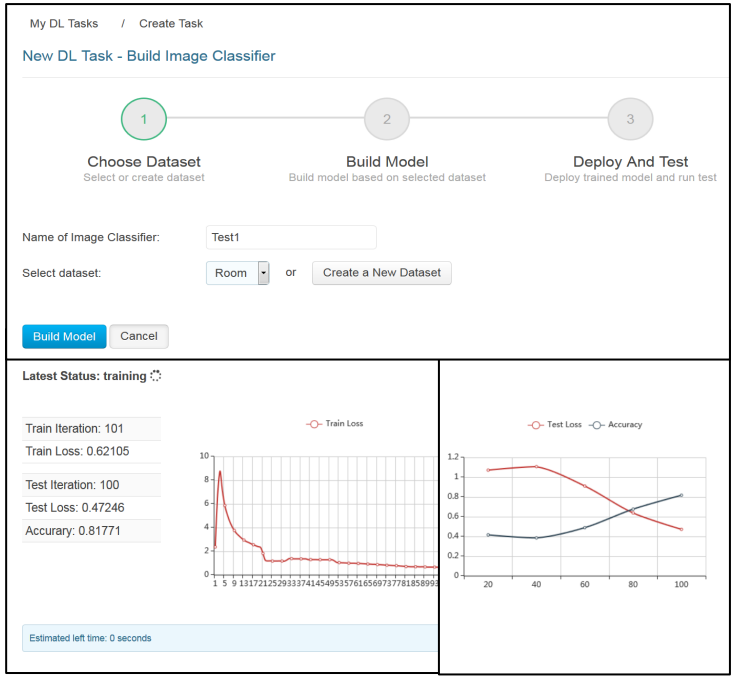
		Deep Learning			ML and DL	Machine Learning
		Power AI Base (WML-CE)	Power AI Enterprise (WML-A)	AI Vision	Watson Studio Local	H2O Driverless AI
Offering	Description	Deep Learning	Deep Learning for the Enterprise	Deep Learning with Video tools	Notebook oriented development environment for ML and DL	Automated 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
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	Coming	Coming	-
	Large Model Support	Yes	Yes	Coming	Coming	-
	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	Coming	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

PowerAI Vision: ”Point-and-Click” AI for Images & Video

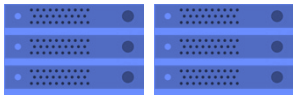
Label Image or Video Data



Auto-Train AI Model

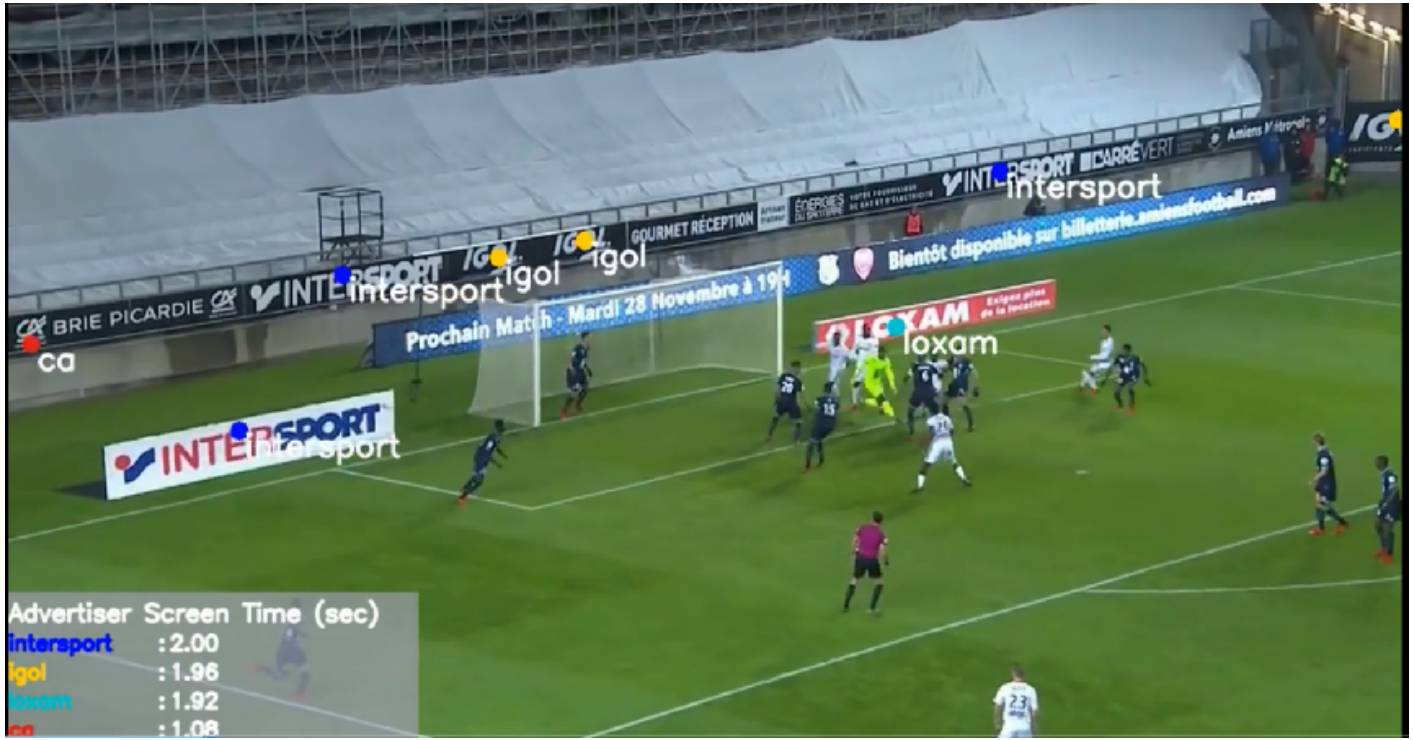


Package & Deploy AI Model

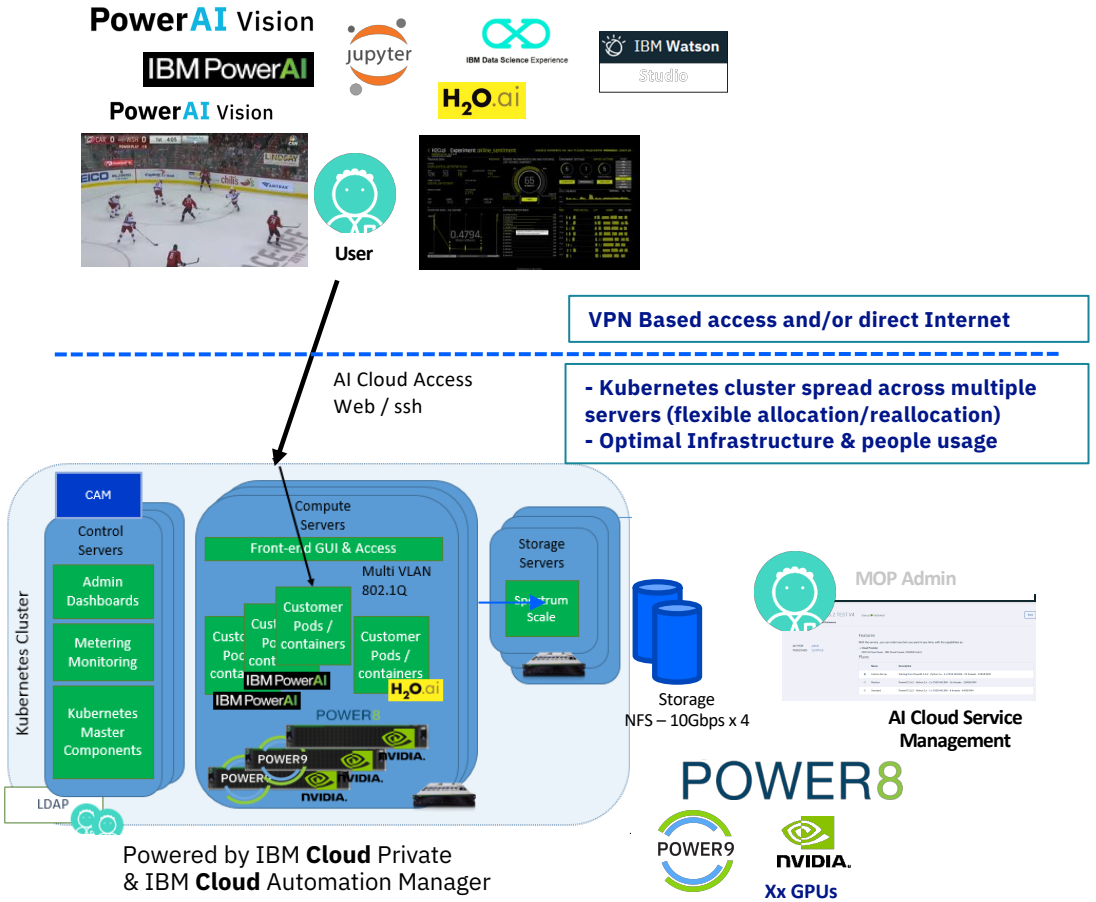


Made with PowerAI Vision – Drone Inspection

Made with PowerAI Vision – Sport / Advertising

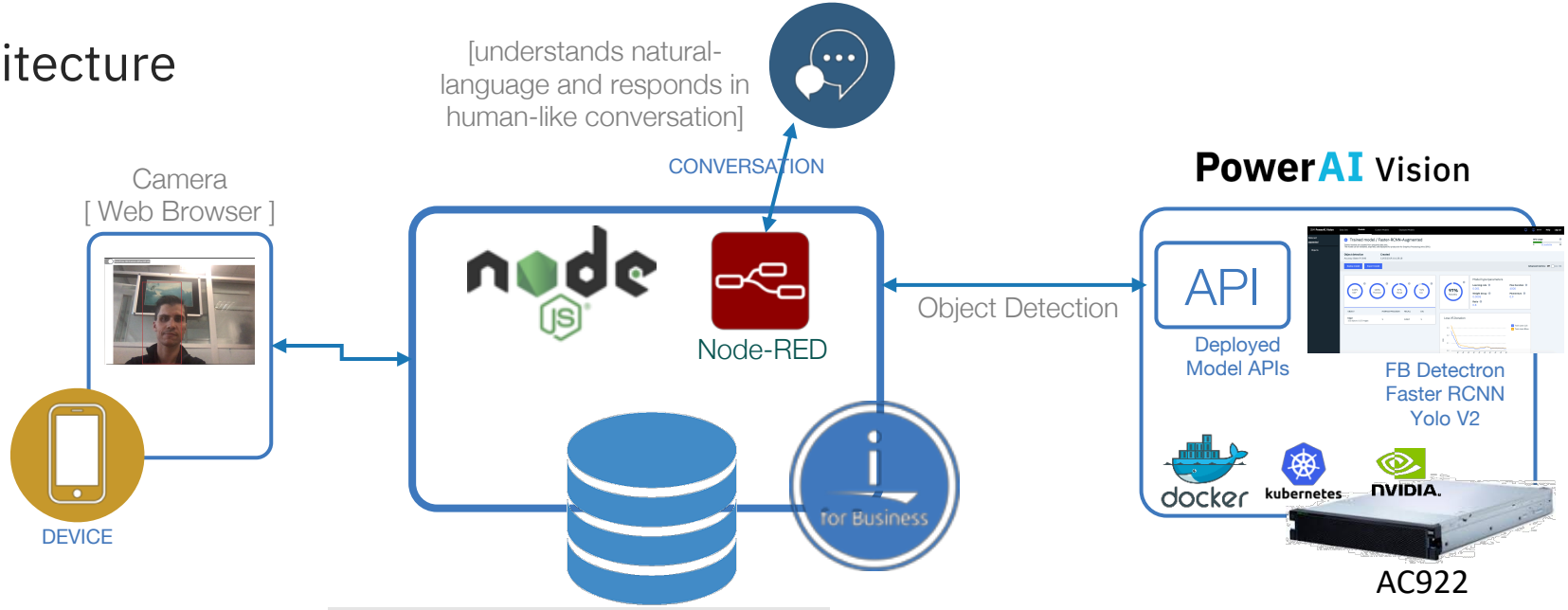


Demo Environment – Private AI Cluster



Computer Vision with PowerAI Vision & IBM i

- Architecture



ID	DETECTEDLABEL	CONFIDENCE	YMAX	XMAX	XMIN	YMIN	A
165	head	0.995372474193573	119.0	111.0	73.0	36.0	en
164	head	0.9906670451164246	119.0	122.0	84.0	62.0	en
163	head	0.9786785244941711	119.0	157.0	123.0	60.0	en
162	head	0.998145580291748	119.0	101.0	53.0	54.0	en
161	head	0.9998226761817932	119.0	95.0	52.0	53.0	en
160	head	0.9996845722198486	119.0	104.0	61.0	53.0	en
159	head	0.9455117583274841	119.0	124.0	83.0	57.0	en
158	head	0.735583484172821	119.0	126.0	87.0	74.0	en
157	head	0.9001550674438477	119.0	119.0	84.0	63.0	en
156	head	0.6558237075805664	119.0	37.0	0.0	52.0	en
155	head	0.9980860948562622	119.0	45.0	3.0	39.0	en
154	head	0.9998077750205994	119.0	103.0	60.0	50.0	en

Computer Vision with PowerAI Vision & IBM i

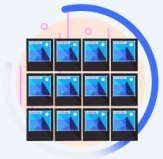
Welcome to IBM PowerAI Vision

Create Dataset

Start by adding images and video files to a data set.

Prepare Data

Label objects or assign categories to images or videos, then use auto labeling to complete the entire data set.



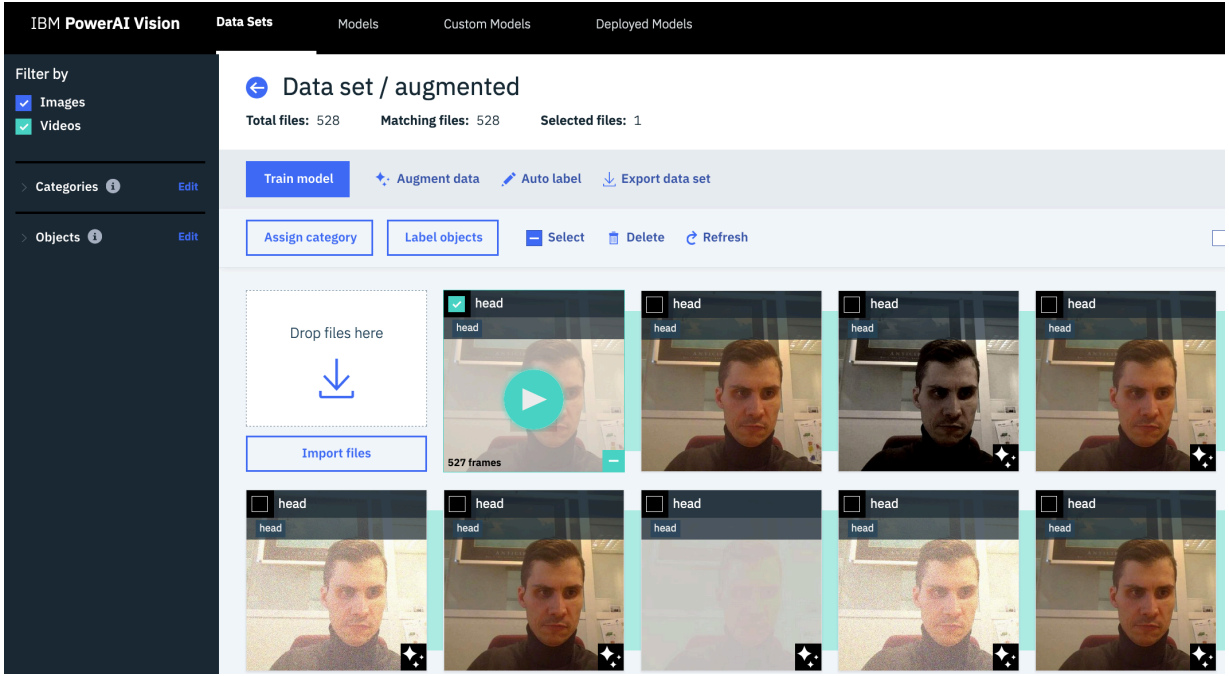
Train Model

Select a few custom options to create your model.

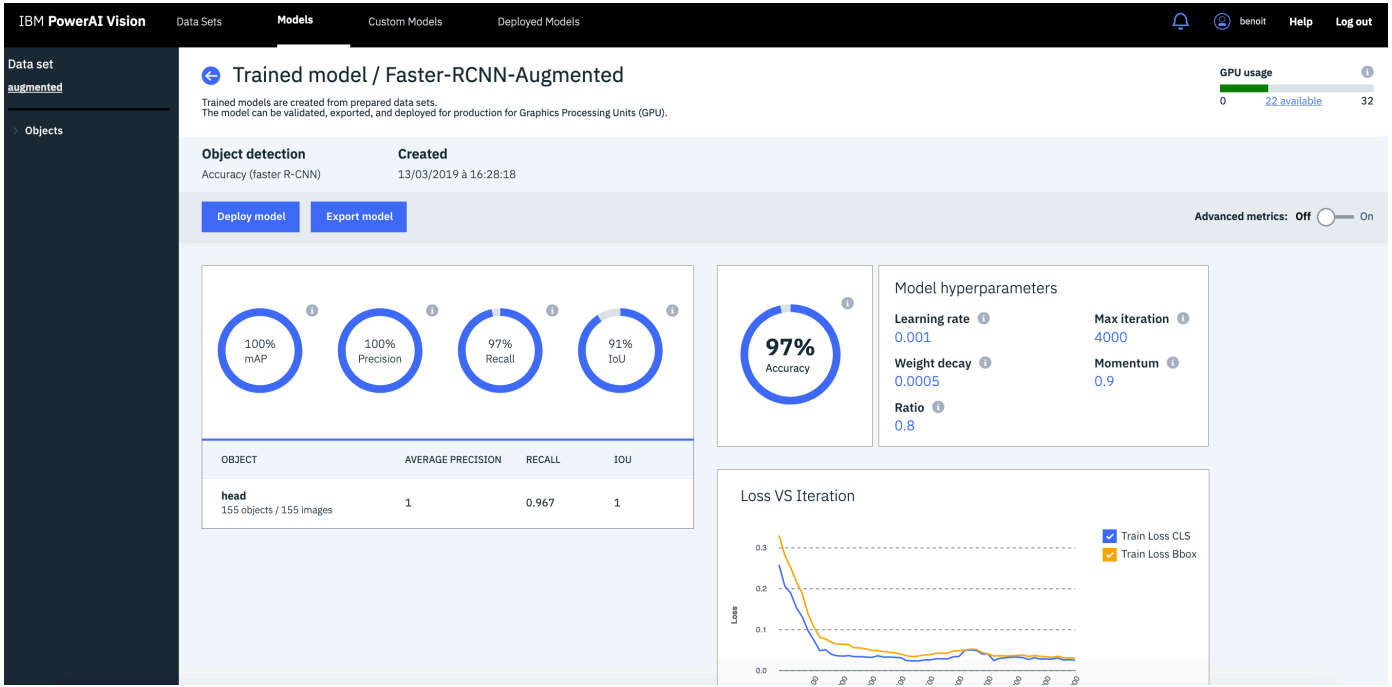
Deploy Model

Deploy the trained model and receive an API link for an inference device.

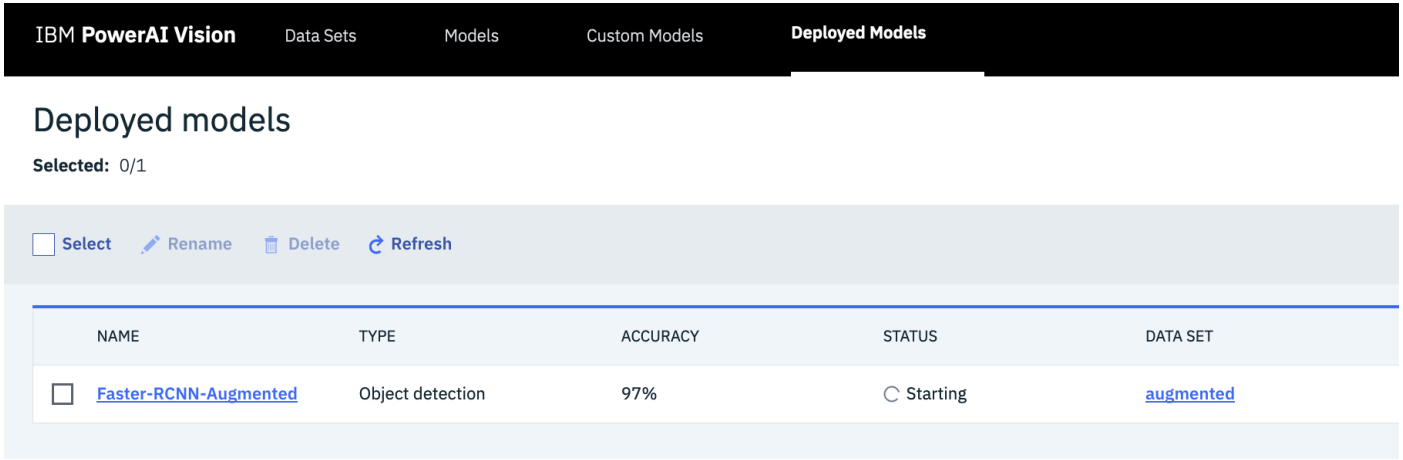
Computer Vision with PowerAI Vision & IBM i



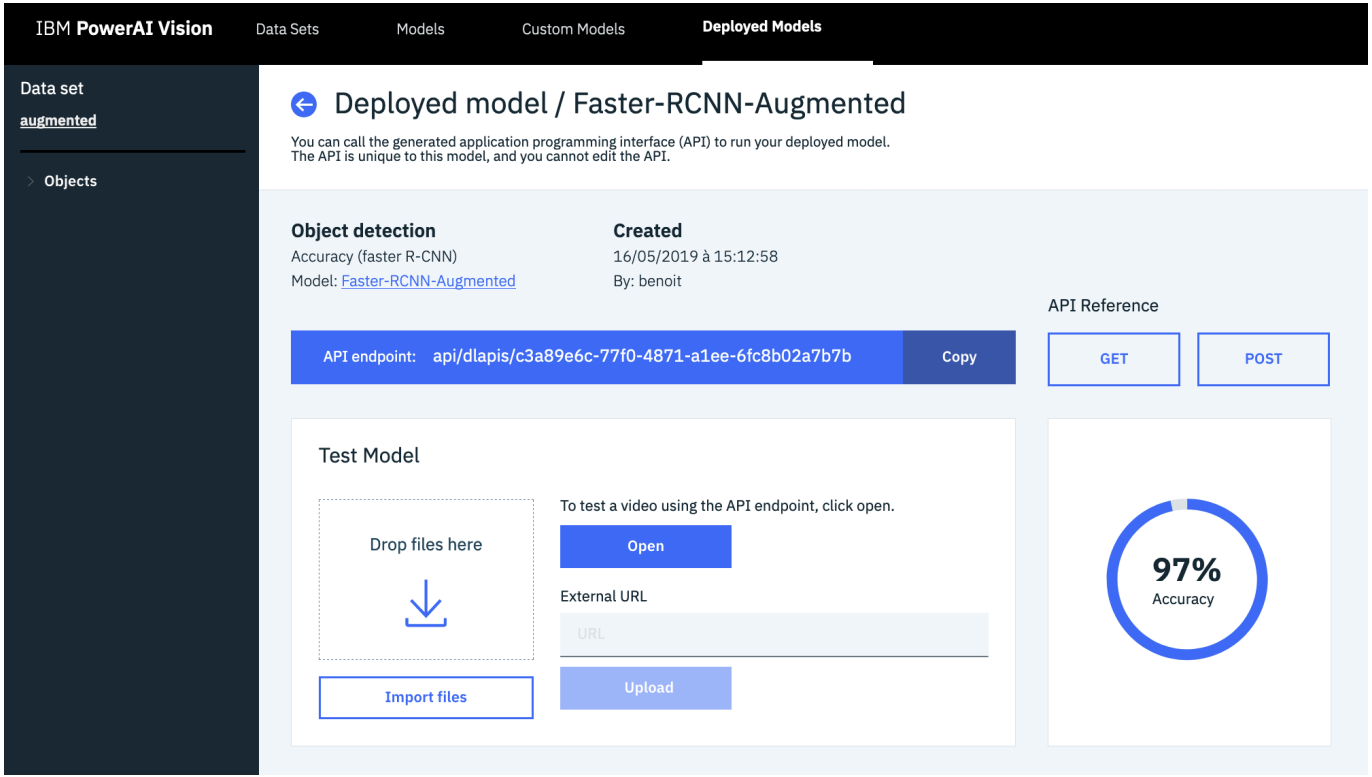
Computer Vision with PowerAI Vision & IBM i



Computer Vision with PowerAI Vision & IBM i



Computer Vision with PowerAI Vision & IBM i



Computer Vision with PowerAI Vision & IBM i

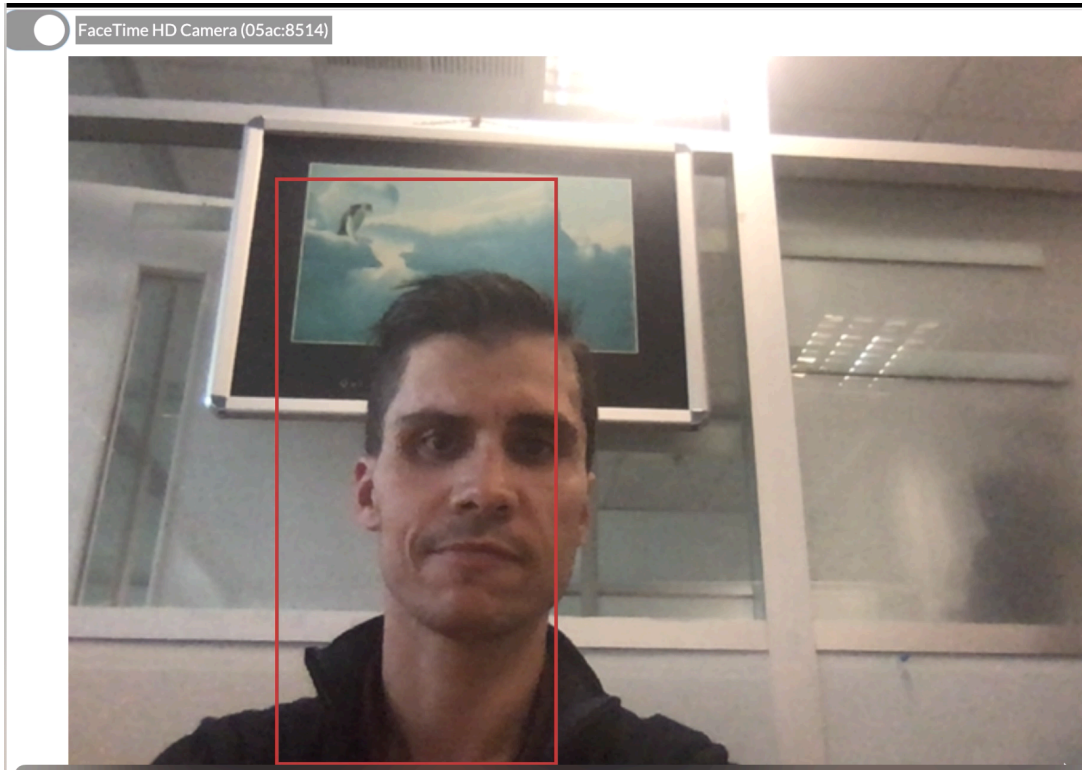
IBM i can now detection objects using an object detection API on premise powered by PowerAI Vision On ppc64le with Nvidia GPUs for inference.

```
[[10:32:25][BENOIT.ICC.LOCAL][~]# system "dsplib aivision"
 5770SS1 V7R3M0  160422                 Display Library                        5/16/19 10:32:43
 Library . . . . . :   AIVISION
 Type  . . . . . :   PROD
 Number of objects . . . . . :          25
 Library ASP number  . . . . . :        1
 Library ASP device  . . . . . :   *SYSBAS
 Library ASP group   . . . . . :   *SYSBAS
 Create authority  . . . . . :   *EXCLUDE
 Text description  . . . . . :   COLLECTION - created by SQL
   Object      Type      Attribute            Size  Description
   QSQJRN0001  *JRNRCV                    1118208  COLLECTION - created by SQL
   QSQJRN0002  *JRNRCV                    1118208  COLLECTION - created by SQL
   QSQJRN0003  *JRNRCV                    1118208  COLLECTION - created by SQL
   QSQJRN0004  *JRNRCV                    1118208  COLLECTION - created by SQL
   QSQJRN      *JRN                         12288  COLLECTION - created by SQL
   DETECTIONS  *FILE     PF                290816
   SYSCHKCST   *FILE     LF                 45056  SQL catalog view
   SYSCOLUMNS  *FILE     LF                135168  SQL catalog view
   SYSCST      *FILE     LF                 86016  SQL catalog view
   SYSCSTCOL   *FILE     LF                 57344  SQL catalog view
   SYSCSTDEP   *FILE     LF                 57344  SQL catalog view
   SYSFIELDS   *FILE     LF                 86016  SQL catalog view
   SYSINDEXES  *FILE     LF                126976  SQL catalog view
   SYSKEYCST   *FILE     LF                 69632  SQL catalog view
   SYSKEYS     *FILE     LF                 86016  SQL catalog view
   SYSPACKAGE  *FILE     LF                 98304  SQL catalog view
   SYSREFCST   *FILE     LF                 77824  SQL catalog view
   SYSTABDEP   *FILE     LF                 77824  SQL catalog view
   SYSTABLES   *FILE     LF                114688  SQL catalog view
   SYSTRIGCOL  *FILE     LF                110592  SQL catalog view
   SYSTRIGDEP  *FILE     LF                114688  SQL catalog view
   SYSTRIGGER  *FILE     LF                167936  SQL catalog view
   SYSTRIGUPD  *FILE     LF                 81920  SQL catalog view
   SYSVIEWDEP  *FILE     LF                 98304  SQL catalog view
   SYSVIEWS    *FILE     LF                 65536  SQL catalog view
                  Total size :            6631424
             *****  E N D  O F  L I S T I N G  *****
```

```
1 select * from aivision.detections order by id desc
```

ID	DETECTEDLABEL	CONFIDENCE	YMAX	XMAX	XMIN	YMIN			
165	head	0.995372474193573	119.0	111.0	73.0	36.			
164	head	0.9906670451164246	119.0	122.0	84.0	62.			
163	head	0.9786785244941711	119.0	157.0	123.0	60.			
162	head	0.998145580291748	119.0	101.0	53.0	54.			
161	head	0.9998226761817932	119.0	95.0	52.0	53.			
160	head	0.9996845722198486	119.0	104.0	61.0	53.0	empty	c3a89e6c-77f0-4871-a1ee-6fc8b02a7b7b	http://powerai-vision-vision-portal:9080/pow
159	head	0.9455117583274841	119.0	124.0	83.0	57.0	empty	c3a89e6c-77f0-4871-a1ee-6fc8b02a7b7b	http://powerai-vision-vision-portal:9080/pow
158	head	0.735583484172821	119.0	126.0	87.0	74.0	empty	c3a89e6c-77f0-4871-a1ee-6fc8b02a7b7b	http://powerai-vision-vision-portal:9080/pow
157	head	0.9001550674438477	119.0	119.0	84.0	63.0	empty	c3a89e6c-77f0-4871-a1ee-6fc8b02a7b7b	http://powerai-vision-vision-portal:9080/pow
156	head	0.6558237075805664	119.0	37.0	0.0	52.0	empty	c3a89e6c-77f0-4871-a1ee-6fc8b02a7b7b	http://powerai-vision-vision-portal:9080/pow
155	head	0.9980860948562622	119.0	45.0	3.0	39.0	empty	c3a89e6c-77f0-4871-a1ee-6fc8b02a7b7b	http://powerai-vision-vision-portal:9080/pow
154	head	0.9998077750205994	119.0	103.0	60.0	50.0	empty	c3a89e6c-77f0-4871-a1ee-6fc8b02a7b7b	http://powerai-vision-vision-portal:9080/pow

Made with PowerAI Vision – Real time detection on IBM i



Visual Recognition Demo on IBM i + PowerAI Vision

for Business

node JS

PowerAI Vision

H2O Driverless AI: How it Works

1 Drag and drop data

Ingest data from cloud, big data and desktop systems

HDFS

SQL

Local

Amazon S3

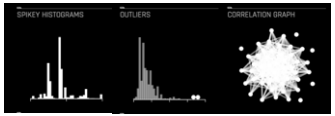
Snowflake

Google BigQuery

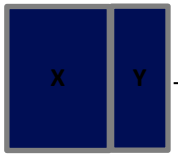
Azure Blog Storage

2 Automatic Visualization

Understand the data shape, outliers, missing values, etc.



Modelling



Dataset

3 Automatic Machine Learning

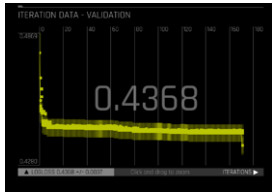
Use best practice model recipes and the power of high performance computing to Iterate across **thousands of possible models** including advanced feature engineering and parameter tuning

Automatic Machine Learning

Model Recipes:
- i.i.d. data
- Time-series
- More on the way

Advanced Feature Engineering + **Algorithm** + **Model Tuning**

Survival of the Fittest



4 Automatic Scoring Pipelines

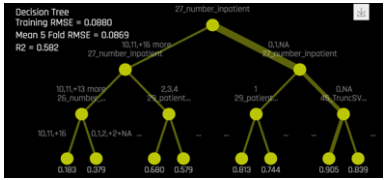
Deploy ultra-low latency Python or Java Automatic Scoring Pipelines that include feature transformations and models.

Model Documentation

Automatic Scoring Pipeline

Deploy Low-latency Scoring to Production

Machine learning Interpretability

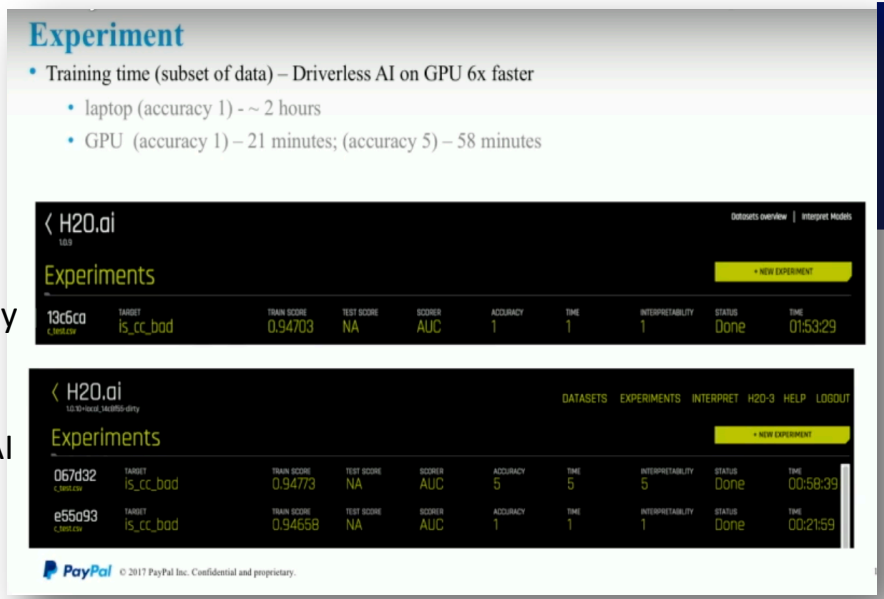


Powered by GPU Acceleration

H2O.ai NVIDIA IBM

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



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

Venkatesh Ramanathan
Senior Data Scientist, PayPal

Leader in Gartner's 2018 Data Science Quadrant

H2O Driverless AI and IBM POWER9 GPU Systems are bringing together the best of breed AI innovation. To handle the increasingly complex workloads of AI you need an integrated system of software and hardware:

- IBM POWER9 supports nearly 2.6x more RAM, 9.5x more I/O bandwidth than comparable systems.
- Nearly 2X the data ingest speed and over 50% faster feature engineering.
- With GPU accelerated machine learning delivering nearly 30X speedup on model building.
- Support for up to 6 V100 GPUs on a single system.

Made with Driverless AI : smarter IBM i apps

Sales + CRM data

Optimized for GPU-Accelerated Power9 Servers

python

Recommendation Engine

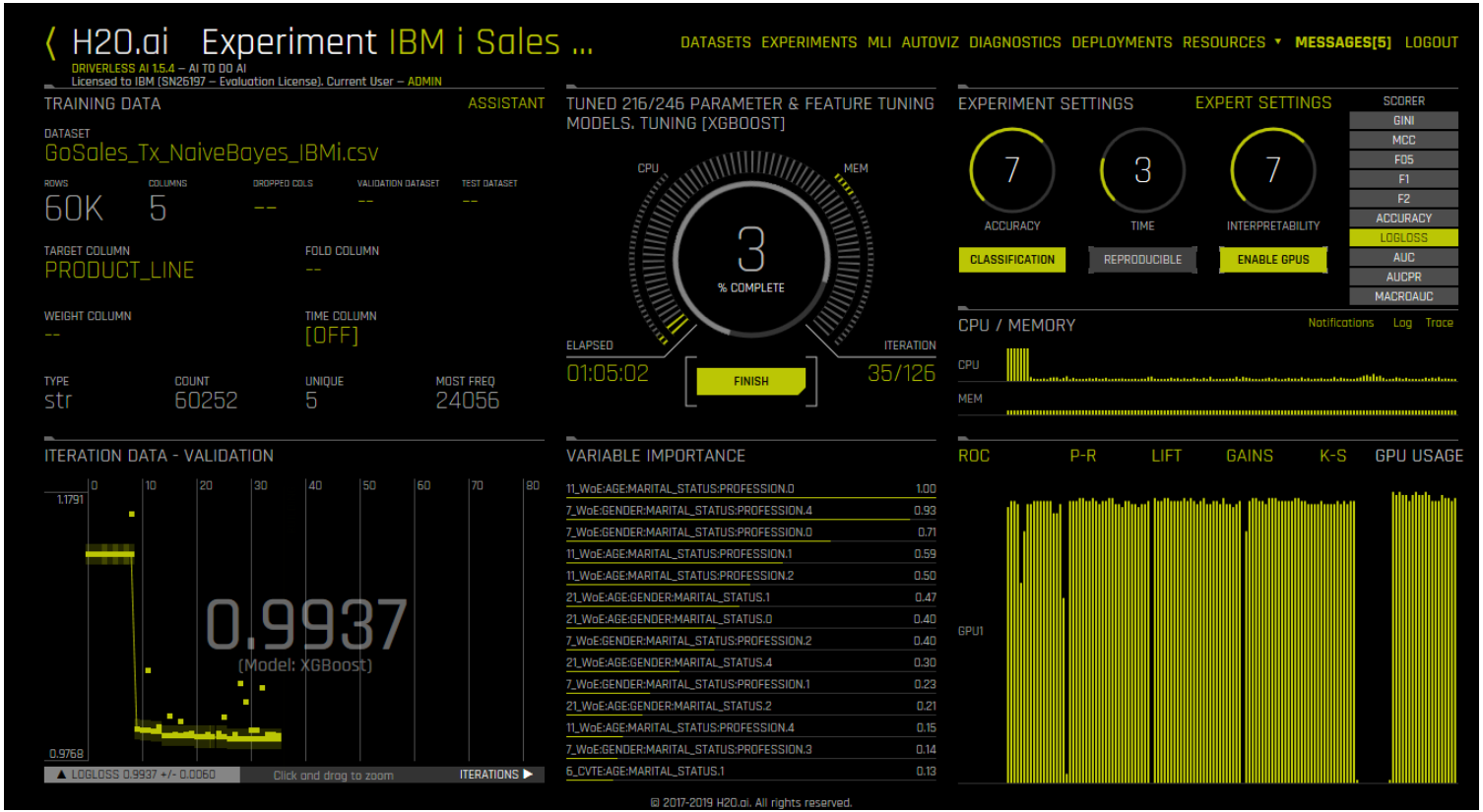
Scoring Pipeline - REST API

python for Business

for Business

```
[14:57:33][sp4i-appserver.power.com][/home/BENOIT/DAI]# uname -a
OS400 sp4i-appserver 2 7 00100006C37T
[14:57:37][sp4i-appserver.power.com][/home/BENOIT/DAI]# ./run_h2o_prediction_from_ibmi.sh M 21.0 Married Retail
This example script demonstrates how to communicate with the Driverless AI Scoring Service via HTTP from IBM i - PASE / Open Source / Any REST Client
The protocol used is JSON-RPC 2.0.
----------------------------------------
Name Type Range Value
----------------------------------------
GENDER object - M
AGE float32 [17.0, 69.0] 21.0
MARITAL_STATUS object - Married
PROFESSION object - Retail
----------------------------------------
Scoring individual rows from IBM i App...
{"jsonrpc": "2.0", "id": 1, "result": [0.7058841586112976, 0.03077976033091545, 0.04269004985690117, 0.022570349276065826, 0.19807571172714233]}
Get the target labels
{"jsonrpc": "2.0", "id": 1, "result": ["Camping Equipment", "Golf Equipment", "Mountaineering Equipment", "Outdoor Protection", "Personal Accessories"]}
```

Made with Driverless AI: smarter IBM i apps



Model Inference from IBM i (Python client)

```
[14:57:33][sp4i-appserver.power.com][/home/BENOIT/DAI]# uname -a
OS400 sp4i-appserver 2 7 00100006C37T
[14:57:37][sp4i-appserver.power.com][/home/BENOIT/DAI]# ./run_h2o_prediction_from_ibmi.sh M 21.0 Married Retail
This example script demonstrates how to communicate with the Driverless AI Scoring Service via HTTP from IBM i - PASE / Open Source / Any REST Client
The protocol used is JSON-RPC 2.0.
----------------------------------------
Name Type Range Value
----------------------------------------
GENDER object - M
AGE float32 [17.0, 69.0] 21.0
MARITAL_STATUS object - Married
PROFESSION object - Retail
----------------------------------------
Scoring individual rows from IBM i App...
{"jsonrpc": "2.0", "id": 1, "result": [0.7058841586112976, 0.03077976033091545, 0.04269004985690117, 0.022570349276065826, 0.19807571172714233]}
Get the target labels
{"jsonrpc": "2.0", "id": 1, "result": ["Camping Equipment", "Golf Equipment", "Mountaineering Equipment", "Outdoor Protection", "Personal Accessories"]}
```

Basic Scikit Learn Modeling vs. Driverless AI demo 1 vs. demo 3

- Same dataset [WA_Fn-UseC_-Telco-Customer-Churn.csv]

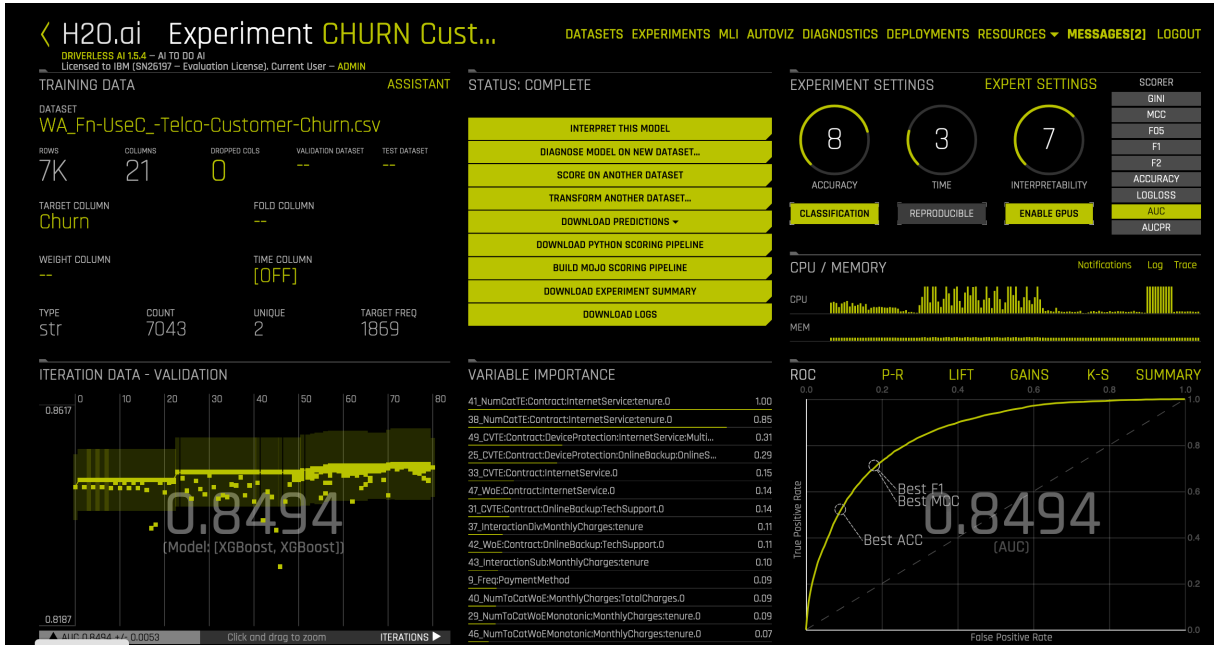
- Accuracy:

0.79 vs. 0.84

- It is just a quick comparison:

Both modeling can be fine-tuned ☺

- Accelerated ML will be able to play with larger datasets
- AutoML (Driverless AI , Watson Studio) will create high quality models, that only experienced data scientists can do




MERCI

Get Started Today

PowerAI Developer Portal
https://developer.ibm.com/linuxonpower/deep-learning-powerai/technology-previews/powerai-vision/

AI Vision Object Detection **Demo**
https://www.youtube.com/watch?v=19vaot75JCY & Jupyter notebook

AI Vision / Public Cloud – Get Started **demo**
https://github.com/IBM/powerai-vision-object-detection

PowerAI **FAQ**
https://developer.ibm.com/linuxonpower/deep-learning-powerai/faq/

PowerAI Vision 1.1.1 **Free trial**
Register for a free 3-day trial of PowerAI Vision

H2o.ai **Driverless AI** (Trial 21 days)
https://www.h2o.ai/products/h2o-driverless-ai/

Presentations, Demo Replays : https://ibm.biz/bma-wiki

Want to know more? Need support ?
Montpellier team & AI Environment for your PoC / Tests: Driverless/PowerAI/AI Vision Remote Access
Contact us : a2roy@fr.ibm.com / benoit.marolleau@fr.ibm.com

ibm.biz/poweraidevelopper