

IBM Client Center Montpellier

Machine Learning & IBM i

Accelerated ML with Driverless AI

DEMONSTRATION



+



Benoit MAROLLEAU - Cloud Architect

IBM Cognitive Systems - Client Center Montpellier, France

benoit.marolleau@fr.ibm.com

<https://ibm.biz/bma-wiki>



About Machine Learning

- ML learns from data, avoids to explicitly program (business) rules.
- ML can establish causation relationships in the Data, when observing a repeating (complex?) phenomenon. The algorithm to build needs examples!
- Business knowledge expertise is necessary in the learning phase to build the algorithm, the model (narrow AI, a model is done for a task)
- Computing resources, libraries & frameworks are necessary
- ML & AI : Domain associated to robotics, philosophy, sociology, mathematics...and more.

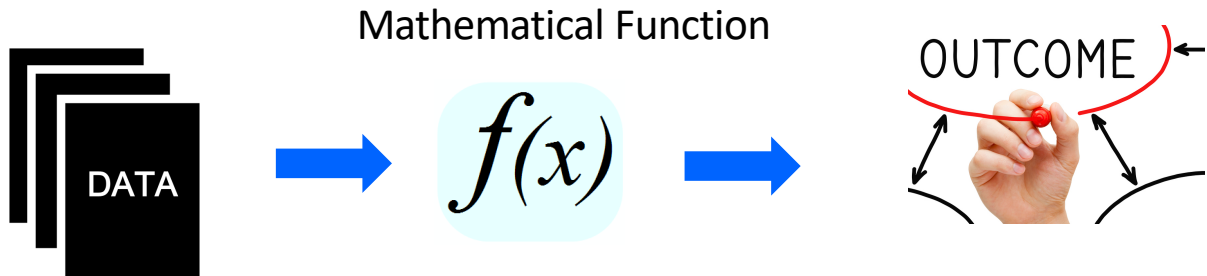
AI cannot be summarized to that, it is an orchestration of ML/DL models (various domains: computer vision, natural language, sensors & feedback...) and traditional techniques (rules engines, expert Systems, and business applications).

About Machine Learning

Why is it a topic of interest?

- Decision making
 - Event / Pattern detection (images, structured data), Classification, customer segmentation
- User Experience (UX)
 - User Interface using image, voice, text, helpdesk, chabot (NLP)
- Augmentation & automation
 - Repetitive tasks, ex: on images, constrained access, complex, autonomous objects.
 - Ex: A doctor could spend hundreds of hours per week to read new studies & publications

About Machine Learning



- Credit card transaction
- Loan application
- MRI image
- House data

- Fraudulent vs. legitimate
- Approve vs. reject
- Tumor benign vs. malignant
- House appraisal value

- **Representing pattern by a mathematical function**
- **Machine learning is just a bunch of math**

About Machine Learning

Objective: Use training data to derive $f(x)$ so that

Minimize (Actual - $f(x)$)

or mathematically

$$\min \quad 1/n \sum_{i=1}^n (y_i - f(x_i))^2$$

- Learning phase : minimize the distance between the current result $f(x)$ and the actual result (supervised learning)
- Volume & Data quality are crucial (Training Dataset). Knowledge of the observed phenomenon, events.
- Each algorithm has **parameters** , adjusted (learned) according to the obtained errors during training.
- The choice of the algorithm is also key, as well as **hyperparameters** initialization.
- Tools & solutions exist to help in these phases.

About Machine Learning

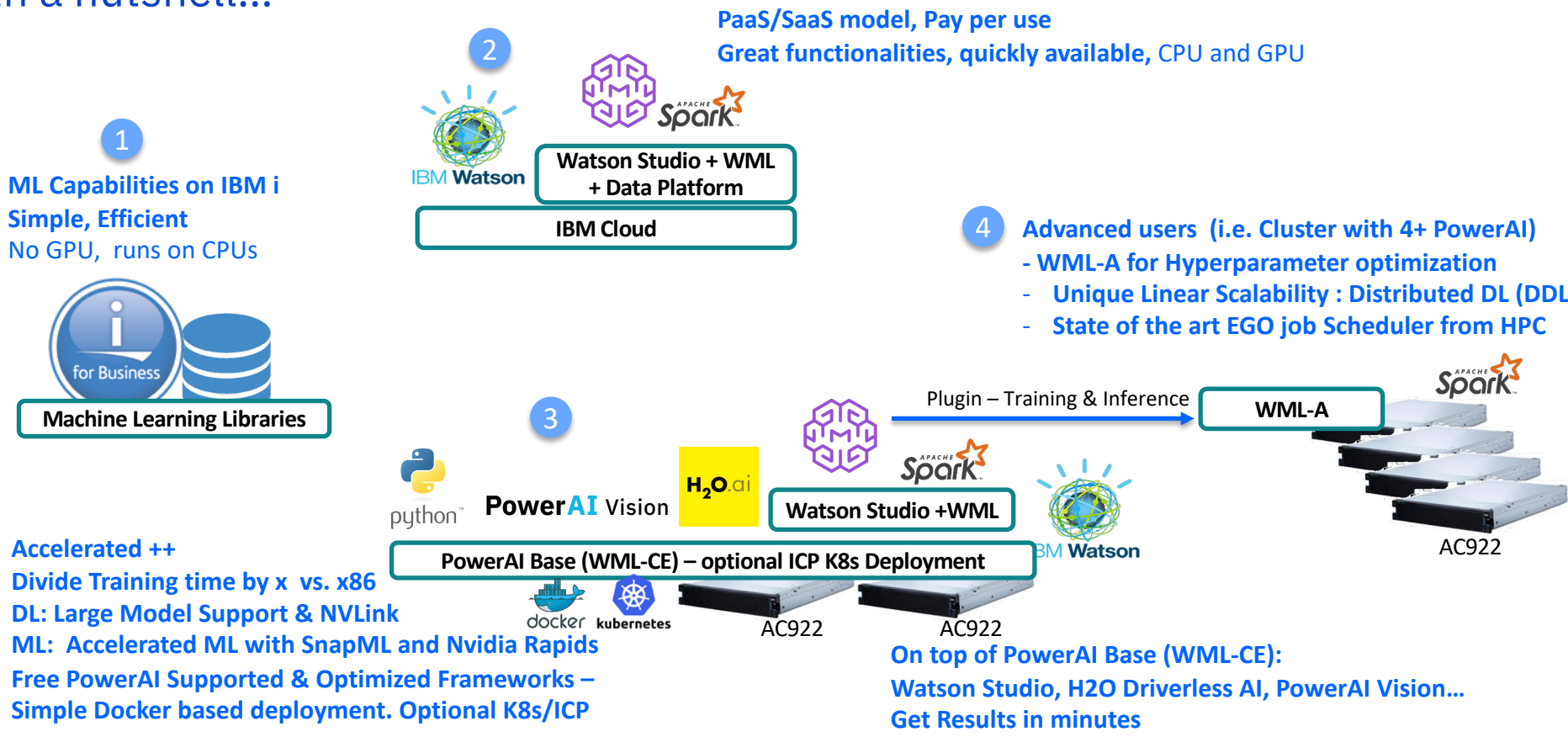
Example – Loan Application

	Feature	Feature	Feature	Feature	Label (=f(X))	
	Loan Requested	Income	Own House	Outstanding Debt	Decision	
Vector X Input	\$20,000	\$100,000	Y	0	Approve	Output
	\$50,000	\$70,000	N	\$20,000	Reject	
	\$5,000	\$150,000	Y	\$10,000	Approve	
	

$$\begin{matrix} x_1 \\ x_2 \\ x_3 \\ \dots \\ x_n \end{matrix} \begin{bmatrix} x_{11} & x_{12} & x_{13} & x_{14} \\ x_{21} & x_{22} & x_{23} & x_{24} \\ x_{31} & x_{32} & x_{33} & x_{34} \\ \dots & \dots & \dots & \dots \\ x_{n1} & x_{n2} & x_{n3} & x_{n4} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \dots \\ y_n \end{bmatrix}$$

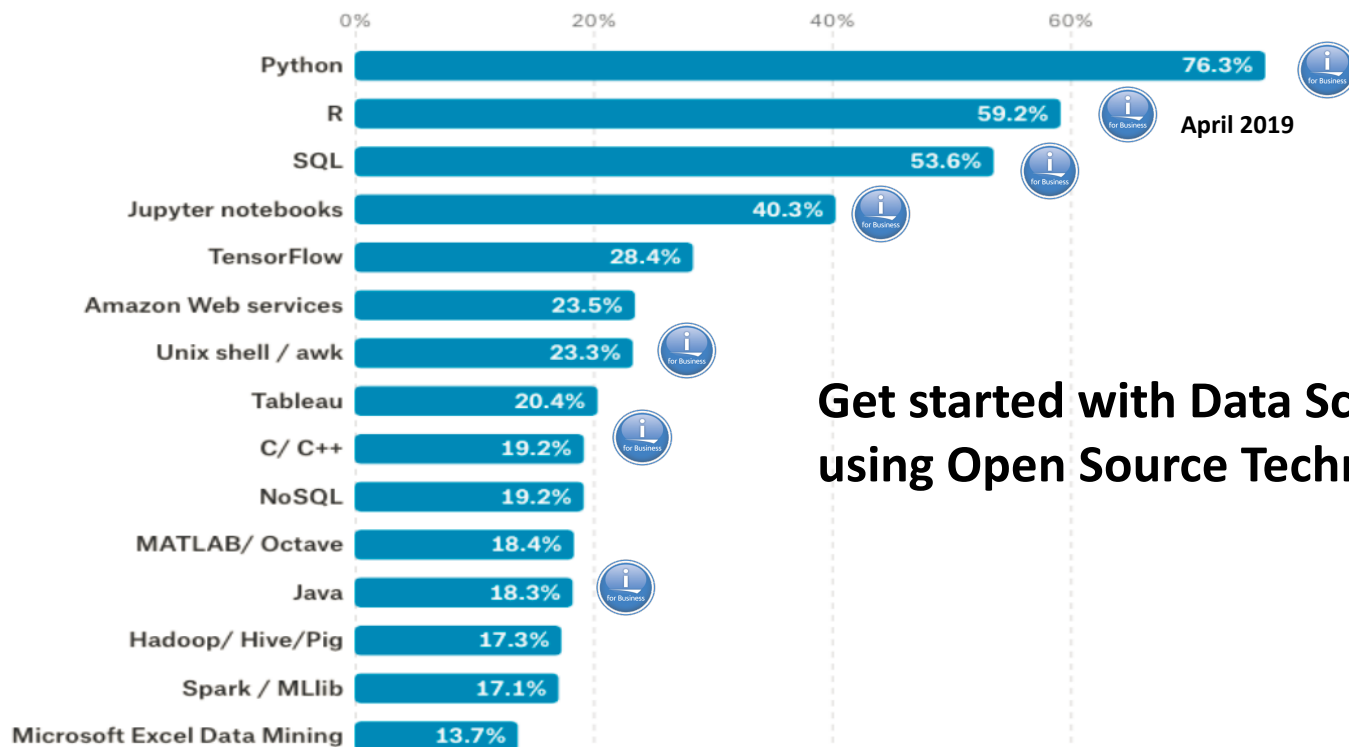
IBM i, Machine Learning & IBM Solutions

In a nutshell...



Data Science tools & technologies

Kaggle 2017 Data Science Tools Survey



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

IBM ML Technologies & IBM i

Scenario 1: Use integrated frameworks , languages on IBM i

Data & Scientific Packages Available

Numpy, Pandas : Data Processing

Scipy, Scikit-Learn

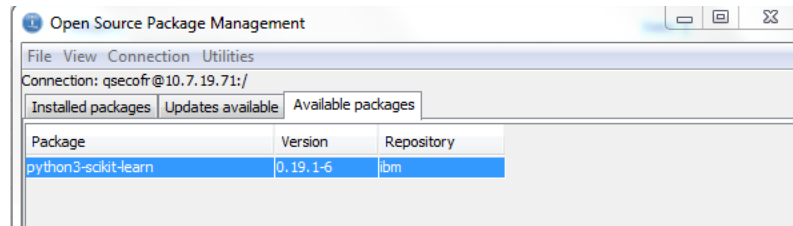
IPython : interactive Python

NLTK : Natural Language Processing

Matplotlib, jupyter : Data Visualization

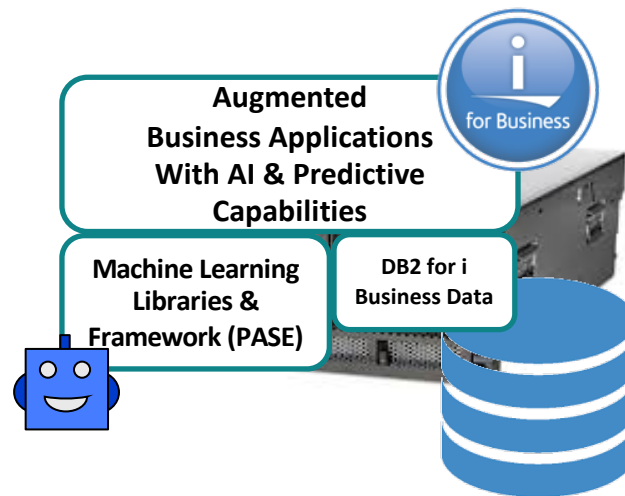
R Language (Interpreter, Runtime)

More to come? 😊



Alternative vs. dataset size, investment on ML etc.

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



Need Acceleration? WML-CE Top 4 Features

Accelerated ML (H2O, Snap ML, Nvidia RAPIDS)

2x to 40x Faster Machine Learning with Snap ML. **H2O Driverless AI automates ML.** Emerging RAPIDS technology. Most enterprise clients use ML today

Large Model Support (LMS)

Our TensorFlow can handle larger models & datasets; leads to higher accuracy

PowerAI Vision

Enables clients without data scientists to start with AI. Bundle with lab services or service provider or workshop

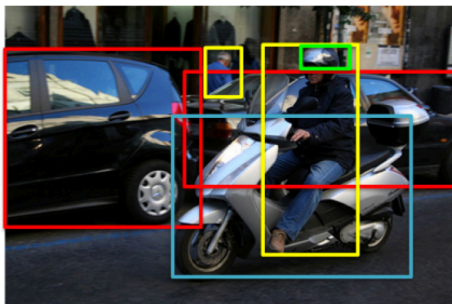
WML Accelerator

Higher Server & GPU Utilization. Target x86 server install base. We can schedule & manage GPU resources better than other software

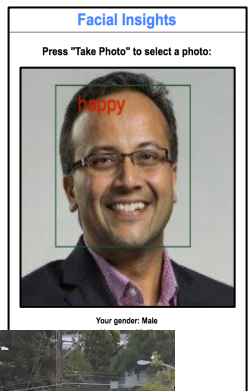
H2O Driverless AI Complements IBM PowerAI Vision



IBM Power AI delivers
Deep Learning for Images



Person
Car
Motorcycle
Helmet



H2O Driverless AI is an
Automatic Machine Learning

Transactional Data: Store Level

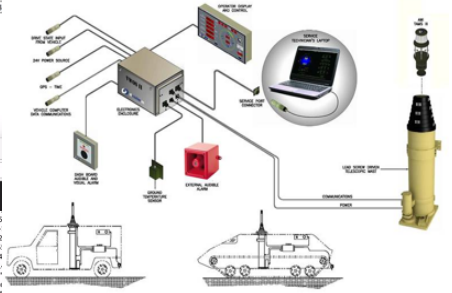
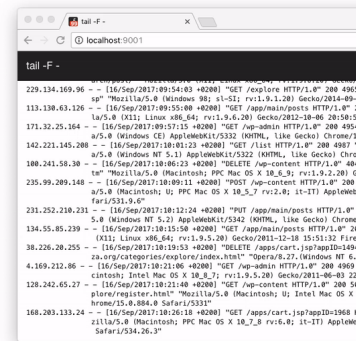
Transactional

Transaction ID	Date	Time	AMOUNT	Card_number
63913045367	11/24/2010	3:20:32 AM	14.47521	*****2323
06781991970	12/4/2010	10:30:10 AM	39.46618	*****6051
0178278302	11/2/2010	1:21:21 AM	24.99964	*****2179
67329082178	10/8/2010	12:29:40 AM	19.21324	*****5826
16835491952	10/17/2010	10:36:12 PM	39.56086	*****9408
08333871154	12/5/2010	5:22:59 PM	33.35401	*****9379
05440123613	11/26/2010	10:36:36 PM	24.23247	*****1816
48906115216	11/3/2010	1:57:13 PM	32.45101	*****2662
14209444560	11/25/2010	7:26:45 PM	23.12291	*****4330
33665080889	10/9/2010	11:09:58 AM	21.93351	*****6533
06738256686	11/23/2010	10:14:36 AM	21.71996	*****4615
08064290443	11/24/2010	1:19:24 AM	15.46741	*****7064
09012945206	11/22/2010	2:02:15 PM	33.14201	*****7140
62116305133	11/23/2010	7:15:13 PM	43.16047	*****9208
17478724264	12/2/2010	12:08:46 PM	40.14018	*****2695
0084400310	10/11/2010	3:10:19 AM	23.70607	*****2980
v2388298971				
w567236458				

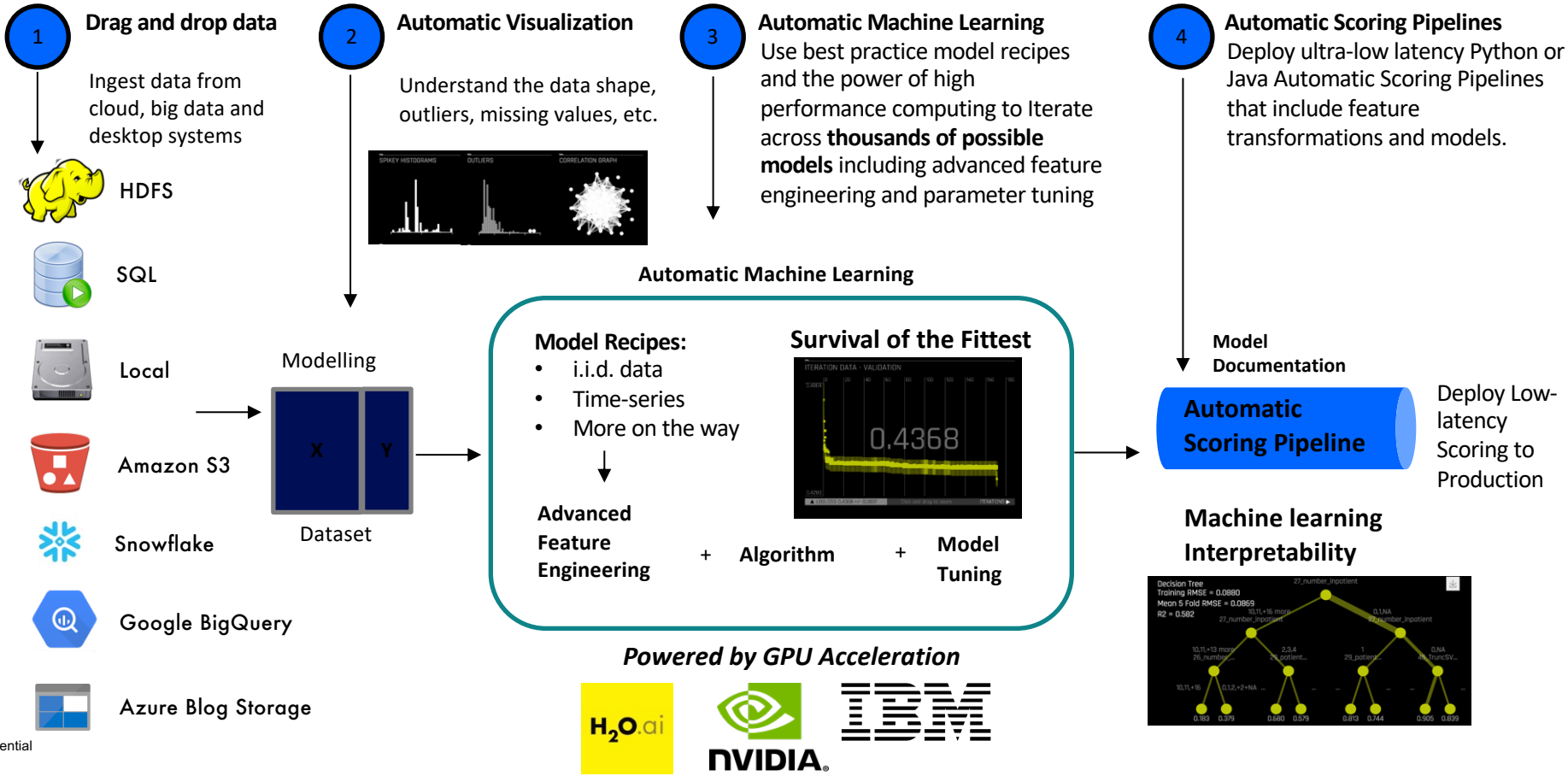
Example: Flat File

Sensors

Log

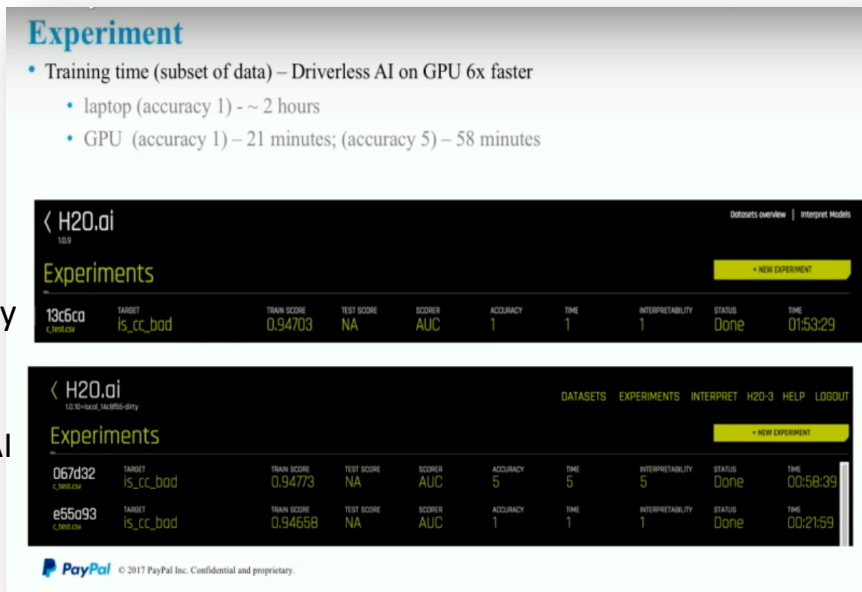


H2O Driverless AI: How it Works



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.

Demonstration

Demo 1 – Introduction to Machine Learning on IBM i: Scikit-Learn demo

Demo 2 - Accelerated Machine Learning with H2O.ai Driverless AI

Demo 3 - Watson Studio – Augment your IBM i App with ML

Demo 4- Deep Learning (Visual Recognition) with PowerAI Vision, IBM i & Node.js

IBM Data Science Technologies & IBM i



- 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



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

- ```

 Terminal Shell Edit View Window Help
Benoit2 — ssh benoit

air-de-benoit:~ Benoit2$ ssh benoit@10.7.19.71
benoit@10.7.19.71's password:

***** IBM i 7.3 Montpellier Client Center *****

#####
#
#
#
#
#
#
#
#####

***** Contact: benoit.marolleau@fr.ibm.com *****

[13:06:15][BENOIT.ICC.LOCAL][~]# cd SKLEARN/
[13:06:18][BENOIT.ICC.LOCAL][~/SKLEARN]# ls -ltr
total 1736
-rw-r--r-- 1 benoit 0 977581 Mar 31 2015 WA_Fn-UseC-Telco-Customer-Churn.csv
drwxr-xr-x 3 benoit 0 8192 May 13 17:57 jupyter
-rw-r--r-- 1 benoit 0 484078 May 14 10:29 SVC_Model_CHURN_Ibmi_V1.joblib
-rw-r--r-- 1 benoit 0 132 May 14 10:51 0_download_Dataset_CSV.sh
-rw-r--r-- 1 benoit 0 393 May 14 11:33 0_load_Dataset_Db2.py
-rw-r--r-- 1 benoit 0 69673 May 14 11:36 customer-churn-prediction.ipynb
drwxr-xr-x 3 benoit 0 8192 May 14 12:28 firstdemo-sckitlearn-ibmi
-rw-r--r-- 1 benoit 0 234 May 14 12:37 README.md
-rw-r--r-- 1 benoit 0 2927 May 14 13:00 1-7_prepare_data_create_train_set.py
-rw-r--r-- 1 benoit 0 1040 May 14 13:03 8-9_feature_engineering.py
-rw-r--r-- 1 benoit 0 457 May 14 13:03 10-11_split_scale.py
-rw-r--r-- 1 benoit 0 164 May 14 13:03 12_build_SVC_ClassifierModel.py
-rw-r--r-- 1 benoit 0 739 May 14 13:04 13_evaluate_SVC_ClassifierModel.py
-rw-r--r-- 1 benoit 0 390 May 14 13:04 14_inference_SVC_ClassifierModel.py
-rw-r--r-- 1 benoit 0 83 May 14 13:04 15_persist_SVC_ClassifierModel.py
-rw-r--r-- 1 benoit 0 78 May 14 13:04 16_load_SVC_ClassifierModel.py
[13:06:29][BENOIT.ICC.LOCAL][~/SKLEARN]#

```

<https://github.com/bmarolleau/firstdemo-scikitlearn-ibmi/>

# IBM Data Science Technologies & IBM i

Demo1

AcsLaunchPad

Untitled\* - Run SQL Scripts - 10.7.19.71(O01dd6f4)

File Edit View Run VisualExplain Monitor Options Connection Tools Help

1 select \* from CHURN.CUSTCHURN2

| F1         | F2     | F3            | F4      | F5         | F6     | F7           | F8               | F9              | F10            |
|------------|--------|---------------|---------|------------|--------|--------------|------------------|-----------------|----------------|
| customerID | gender | SeniorCitizen | Partner | Dependents | tenure | PhoneService | MultipleLines    | InternetService | OnlineSecurity |
| 7590-VHVEG | Female | 0             | Yes     | No         | 1      | No           | No phone service | DSL             | No             |
| 5575-GNVDE | Male   | 0             | No      | No         | 34     | Yes          | No               | DSL             | Yes            |
| 3668-QPYBK | Male   | 0             | No      | No         | 2      | Yes          | No               | DSL             | Yes            |
| 7795-CFOCW | Male   | 0             | No      | No         | 45     | No           | No phone service | DSL             | Yes            |
| 9237-HQITU | Female | 0             | No      | No         | 2      | Yes          | No               | Fiber optic     | No             |
| 9305-CDSKC | Female | 0             | No      | No         | 8      | Yes          | Yes              | Fiber optic     | No             |
| 1452-KIOVK | Male   | 0             | No      | Yes        | 22     | Yes          | Yes              | Fiber optic     | No             |
| 6713-OKOMC | Female | 0             | No      | No         | 10     | No           | No phone service | DSL             | Yes            |
| 7892-POOKP | Female | 0             | Yes     | No         | 28     | Yes          | Yes              | Fiber optic     | No             |
| 6388-TABGU | Male   | 0             | No      | Yes        | 62     | Yes          | No               | DSL             | Yes            |
| 9763-GRSKD | Male   | 0             | Yes     | Yes        | 13     | Yes          | No               | DSL             | Yes            |
| 7469-LKBCI | Male   | 0             | No      | No         | 16     | Yes          | No               | No              | No internet se |
| 8091-TTVAX | Male   | 0             | Yes     | No         | 58     | Yes          | Yes              | Fiber optic     | No             |
| 0280-XJGEX | Male   | 0             | No      | No         | 49     | Yes          | Yes              | Fiber optic     | No             |
| 5129-JLPIS | Male   | 0             | No      | No         | 25     | Yes          | No               | Fiber optic     | Yes            |
| 3655-SNOYZ | Female | 0             | Yes     | Yes        | 69     | Yes          | Yes              | Fiber optic     | Yes            |

100 rows displayed (more data available).

Messages Global Variables and Special Registers select \* from CHURN.CUSTCHURN2

Connected to relational database O01dd6f4 on 10.7.19.71 as QSECOFR - 136910/QUSER/QZDASOINIT using JDBC configuration 'Default'.

- System Configurations
- 5250 Session Manager
- HMC Probe Utility

F3=Exit F12=Cancel

Start Transfer Stop Transfer Properties

Rows transferred: 7,044

# IBM Data Science Technologies & IBM i

jupyter Churn-IBMi (autosaved) Logout

File Edit View Insert Cell Kernel Help Trusted Python 3

Run

## Machine Learning on IBM i with Scikit Learn, pandas (0.22) and Db2 for i

Original notebook: <https://github.com/IBM/customer-churn-prediction/blob/master/notebooks/customer-churn-prediction.ipynb>

```
In [6]: %pwd
Out[6]: '/home/BENOIT/firstdemo-scikitlearn-ibmi'
```

```
In [9]: # Download the original dataset if necessary
#!wget https://community.watsonanalytics.com/wp-content/uploads/2015/03/WA_Fn-UseC_-Telco-Customer-Churn.csv --no-check-certificate
```

### Load Dataframe from Db2

```
In [10]: import ibm_db_dbi as dbi
import pandas as pd

Telco Churn dataset from a CRM datasource - here Db2 for i in CHURN/CUSTCHURN2 7044 records

sql="SELECT * from CHURN.CUSTCHURN3"
pd.set_option('display.max_columns', 30)
try:
 conn = dbi.connect()
 df_fromdb2 = pd.read_sql(sql, conn)
 #print(df_fromdb2)
 df_fromdb2.info()
```

# Demonstration

Demo 1 – Introduction to Machine Learning on IBM i: Scikit-Learn demo

Demo 2 - Accelerated Machine Learning with H2O.ai Driverless AI

Demo 3 - Watson Studio – Augment your IBM i App with ML

Demo 4- Deep Learning (Visual Recognition) with PowerAI Vision, IBM i & Node.js

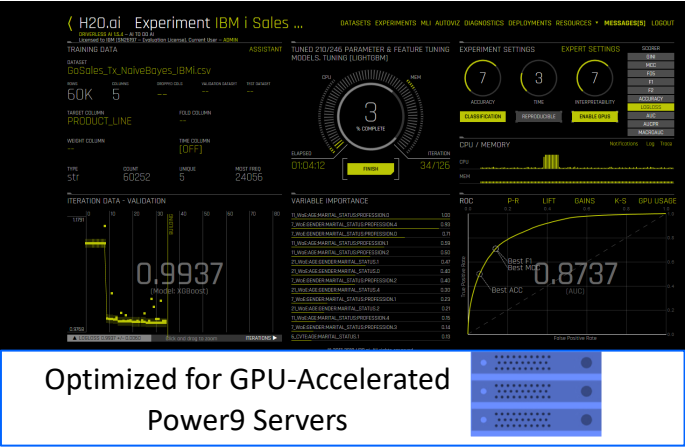
# Smarter IBM i apps made easy with Driverless AI

## CRM & Customer Churn scenario

# Summary

Import & Visualize Data  
Create & Test Models

Extract  
CRM data  
(or JDBC)



Model Export



Recommendation Engine  
Scoring Pipeline

Augment IBM i Applications  
Model Inference

Monitor Models  
& Iterate



Business Database

```
[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"]}
```

Business Application

- Terminal window showing a remote session to a server. The terminal displays the login process, directory listing, and execution of a SQL query. The query results are shown in a table with columns F1 through F10. The table lists customer information including ID, gender, senior citizen status, partner status, dependents, tenure, phone service, multiple lines, internet service, and online security. The results show 100 rows displayed.

```
air-de-benoit:~ Benoit2$ ssh benoit@10.7.19.71
benoit@10.7.19.71's password:

***** IBM i 7.3 Montpellier Client Center *****

#####
#
#
#
#
#
#####

***** Contact: benoit.marolleau@fr.ibm.com *****

[[13:06:15][BENOIT.ICC.LOCAL][~]# cd SKLEARN/
[[13:06:18][BENOIT.ICC.LOCAL][~/SKLEARN]# ls -ltr
total 1736
-rw-r--r-- 1 benoit 0 977501 Mar 31 2015 WA_Fn-UseC-_Telco-Customer-Churn.csv
drwxr-sr-x 3 benoit 0 8192 May 13 17:57 jupyter
-rw-r--r-- 1 benoit 0 484078 May 14 10:29 SVC_Model_CHURN_IBMi_V1.joblib
-rw-r--r-- 1 benoit 0 132 May 14 10:51 0_download_Dataset_CSV.sh
-rw-r--r-- 1 AcsLaunchPad

Unittied* - Run SQL Scripts - 10.7.19.71[001dd6f4]
File Edit View Run Visual Explain Monitor Options Connection Tools Help
1 select * from CHURN.CUSTCHURN2

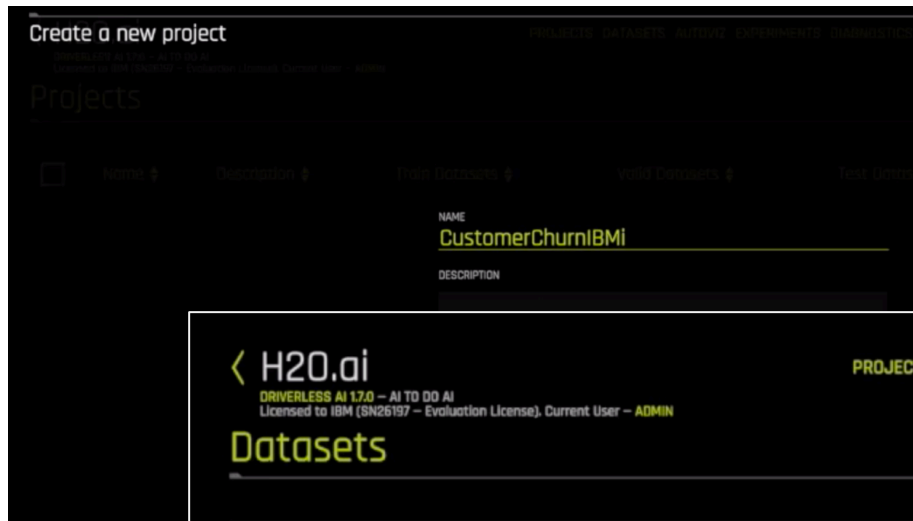
F1 F2 F3 F4 F5 F6 F7 F8 F9 F10
customerID gender SeniorCitizen Partner Dependents tenure PhoneService MultipleLines InternetService OnlineSecurity
7590-VHVEG Female 0 Yes No 1 No No phone service DSL No
5575-QWDE Male 0 No No 34 Yes No DSL Yes
3669-QPYBK Male 0 No No 22 Yes Fiber optic No
7795-CFCOW Male 0 No No 45 No No phone service DSL Yes
9237-HQITU Female 0 No No 2 Yes No Fiber optic No
9385-CDSKC Female 0 No No 8 Yes Yes Fiber optic No
1452-KIDUK Male 0 No Yes 22 Yes Fiber optic No
6713-KOMCK Female 0 No No 10 No No phone service DSL Yes
7892-POOKP Female 0 Yes No 28 Yes Yes Fiber optic No
6388-TABGU Male 0 No Yes 62 Yes No DSL Yes
9763-GRSKD Male 0 Yes Yes 13 Yes No DSL Yes
7469-LKBCI Male 0 No No 16 Yes No No No internet service
8891-TTVAX Male 0 Yes No 58 Yes Yes Fiber optic No
8298-XJGEX Male 0 No No 49 Yes Yes Fiber optic No
5129-JLFTS Male 0 No No 25 Yes No Fiber optic Yes
3655-SNOYZ Female 0 Yes Yes 69 Yes Yes Fiber optic Yes

100 rows displayed (more data available).
Messages Global Variables and Special Registers select * from CHURN.CUSTCHURN2

Connected to relational database 001dd6f4 on 10.7.19.71 as QSECOFR - 136910/QUSER/QZDASOINIT using JDBC configuration 'Default'.
System Configurations
5250 Session Manager
HMC Probe Utility
Start Transfer Stop Transfer Properties
F3=Exit F12=Cancel
Rows transferred: 7,044
```

# Dataset Import in Driverless AI

2



**Drag and drop data – Churn.csv**  
Ingest data from Db2 for i

This screenshot shows the H2O.ai 'Datasets' page. The page header includes the H2O.ai logo, version 'DRIVERLESS AI 17.0 - AI TO DO AI', and user information 'Licensed to IBM (SN26197 - Evaluation License), Current User - ADMIN'. The 'Datasets' section has a '+ ADD DATASET (OR DRAG & DROP)' button. Below is a table listing datasets, with the first row highlighted by a red box.

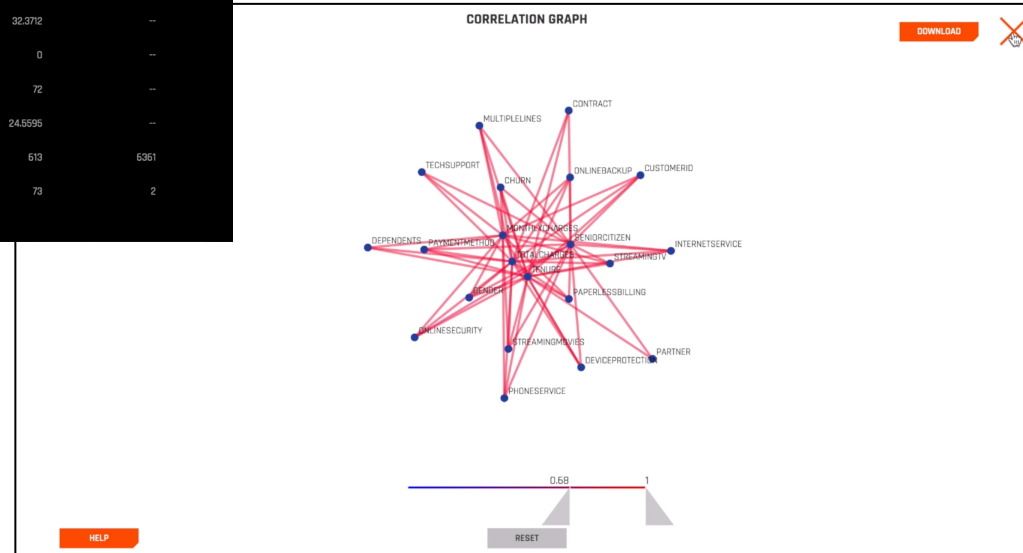
| <input type="checkbox"/> | Name                                 | Path                                         | Size | Rows  | Columns | Status              |
|--------------------------|--------------------------------------|----------------------------------------------|------|-------|---------|---------------------|
| <input type="checkbox"/> | Churn.csv                            | ./tmp/ae192e76-c425-11e9-80bd-325492eff8...  | 1MB  | 7043  | 21      | [Click for Actions] |
| <input type="checkbox"/> | UCI_Credit_Card.csv                  | ./tmp/cohutuno/UCI_Credit_Card.csv.156222... | 4MB  | 30000 | 25      | [Click for Actions] |
| <input type="checkbox"/> | GoSales_Tx_NaiveBayes.csv            | ./tmp/moceneko/GoSales_Tx_NaiveBayes.cs...   | 3MB  | 60252 | 5       | [Click for Actions] |
| <input type="checkbox"/> | WA_Fn-UseC_-Telco-Customer-Churn.csv | ./tmp/sigolaga/WA_Fn-UseC_-Telco-Custom...   | 1MB  | 7043  | 21      | [Click for Actions] |

# Data Visualization on Driverless AI

2



**Automatic Visualization**  
Understand the data shape, outliers, missing values, etc.



# Create new model

2

**H2O.ai Experiment**  
DRIVERLESS AI 17.0 - AI TO DO AI  
Licensed to IBM (SH26197 - Evaluation License). Current User - ADMIN

PROJECTS DATASETS AUTOVIZ EXPERIMENTS DIAGNOSTICS MLJ DEPLOYMENTS RESOURCE

TRAINING DATA ASSISTANT

DATASET: Churn.csv

ROWS: 7K COLUMNS: 21 DROPPED COLS: -- VALIDATION DATASET: -- TEST DATASET: --

TARGET COLUMN: CHURN FOLD COLUMN: --

WEIGHT COLUMN: -- TIME COLUMN: [OFF]

TYPE: STR COUNT: 7043 UNIQUE: 2 TARGET FREQ: 1869

EXPERIMENT SETTINGS EXPERT SETTINGS

ACCURACY: 6 TIME: 3 INTERPRETABILITY: 7 SCORER: AUG

CLASSIFICATION REPRODUCIBLE GPU ENABLED

LAUNCH EXPERIMENT

What do these settings mean?

ACCURACY: 6

- Training data size: 7,043 rows, 21 cols
- Feature evolution: [XGBoostGBM, LightGBM], 3-fold CV, 2 reps
- Final pipeline: Ensemble (6 models), 3-fold CV

TIME: 3

- Feature evolution: 4 individuals, up to 62 iterations
- Early stopping: After 5 iterations of no improvement

INTERPRETABILITY: 7

- Feature pre-pruning strategy: FS
- XGBoost Monotonicity constraints: enabled
- Feature engineering search space (where applicable): (Original, WeightOfEvidence, IsHoliday, Dates, CVTargetEncode, Interactions, NumToCatVols, TextLinModel, Text, NumToCatWoEMonotonic, NumToCatTE, CVCatNumEncode, NumCatTE, Frequent)

[XGBoostGBM, LightGBM] models to train:

- Model and feature tuning: 192
- Feature evolution: 384
- Final pipeline: 6

Estimated runtime: minutes

Whether to enable GPU(s). Only supported on machine accelerators. Preferred are multi-GPU boxes with Pascal

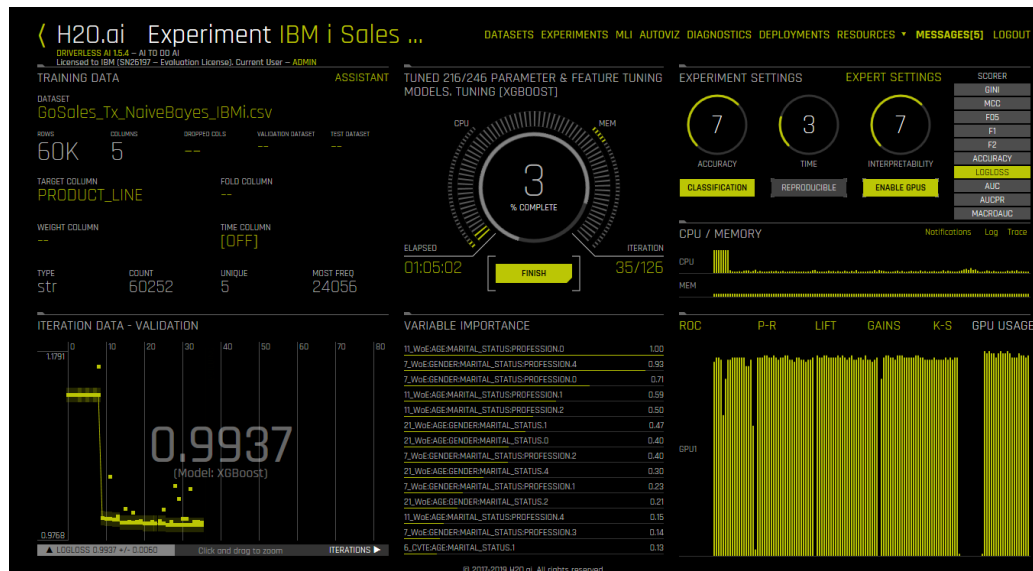
## Automatic Machine Learning

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

*Target Column : Churn (YES/NO) for binary classification*

# Create new model

3



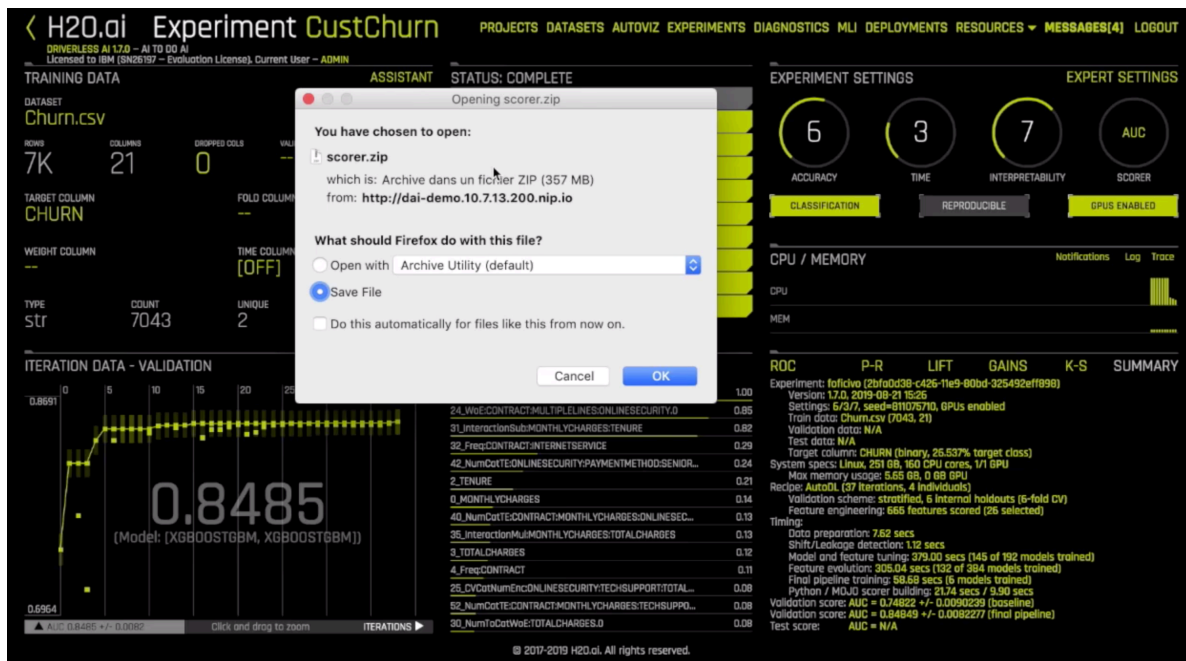
## Automatic Machine Learning

- supports nearly 2.6x more RAIBM POWER9 M, 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

*Experiment running –GPU Accelerated AutoML on POWER9*

# Export Scoring Pipeline

3



Scoring Pipeline export

## Automatic Scoring Pipelines

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

Here, the Python Scoring Pipeline “scorer.zip” to be deployed on the inference system

# Deploy Low-latency Scoring to Production on IBM i

```

Benoit2 — ssh benoit@10.7.19.71 — 162x41
air-de-benoit:~ Benoit2$
air-de-benoit:~ Benoit2$ ssh benoit@10.7.19.71
benoit@10.7.19.71's password:

***** IBM i 7.3 Montpellier Client Center *****

#####
#
#
#
#
#

***** Contact: benoit.marolleau@fr.ibm.com *****

[[11:50:13][BENOIT.ICC.LOCAL][~]# cd ~/DAI/CustomChurn
[[11:50:23][BENOIT.ICC.LOCAL][~/DAI/CustomChurn]# ls
scorer.zip scoring-pipeline
[[11:50:24][BENOIT.ICC.LOCAL][~/DAI/CustomChurn]# cd scoring-pipeline/
[[11:50:40][BENOIT.ICC.LOCAL][~/DAI/CustomChurn/scoring-pipeline]# ls
README.txt client_requirements.txt
common-functions.sh datatable-0.7.0.dev479-cp36m-linux_ppc64le.whl
environment.yml example.py
example_client.py h2oai_experiment_summary_higikobe.zip
h2oai_experiment_summary_higikobe.zip
http_server.py http_server_requirements.txt
requirements.txt run_http_client.sh
run_http_client.sh run_tcp_client.sh
scoring.thrift tcp_server.py
tcp_server.py thrift_post_processing.py
[[11:50:42][BENOIT.ICC.LOCAL][~/DAI/CustomChurn/scoring-pipeline]#

[[11:50:42][BENOIT.ICC.LOCAL][~/DAI/CustomChurn/scoring-pipeline]# ./predictCustomerChurn.sh
Scoring individual rows on IBM i...
row: {
 Partner: No,
 tenure: 1.0,
 PhoneService: No,
 MultipleLines: Yes,
 InternetService: Fiber optic,
 OnlineSecurity: No,
 OnlineBackup: No internet service,
 DeviceProtection: No,
 TechSupport: No internet service,
 StreamingMovies: No internet service,
 Contract: One year,
 PaperlessBilling: Yes,
 PaymentMethod: Mailed check,
 MonthlyCharges: 20.850000381469727,
 TotalCharges: 70.1500015258789
}

{"jsonrpc": "2.0", "id": 1, "result": [0.8188602924346924, 0.18113969266414642]}
Get the target labels
{"jsonrpc": "2.0", "id": 1, "result": ["No", "Yes"]}

```

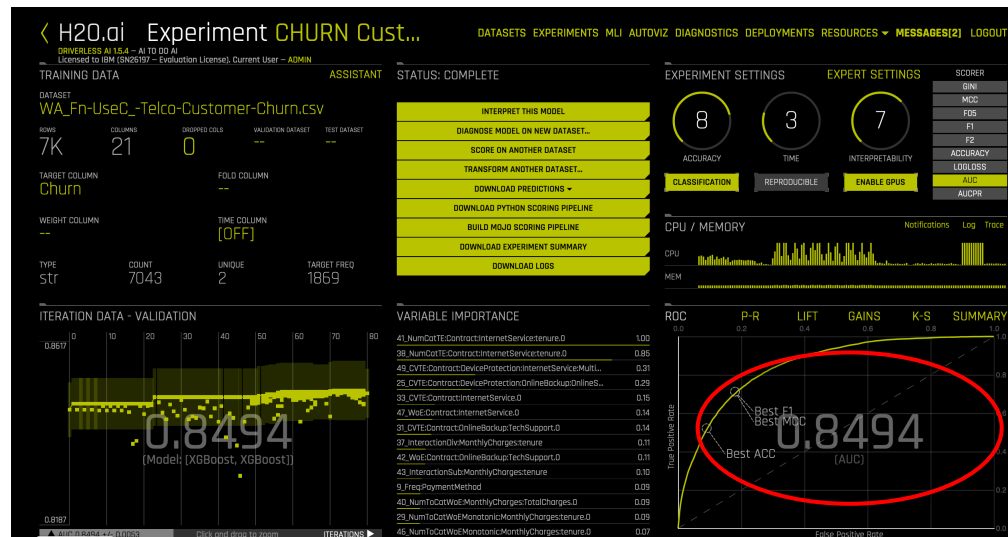
- Invoke low latency scoring pipelines (real time, batch)
- Augment any IBM i apps (ILE, Java, Python,...) with Machine Learning models.

Model Inference with Live CRM data in Production (Db2)

Scoring Result  
Customer classified in category Churn = NO , 0.81% Confidence

# Manual Modeling vs. Driverless AI Accelerated Auto ML

- Customer Churn Dataset [WA\\_Fn-UseC\\_-Telco-Customer-Churn.csv](#)  
*Customer churn is when an existing customer, user, player, subscriber or any kind of return client stops doing business or ends the relationship with a company.*
- Model Accuracy: 0.79 vs. **0.84**
- Time to market & effort: Days vs. Minutes with DAI  
AutoML w/ Feature Engineering, etc.  
What's the business impact?
- It is just a quick comparison:  
Both modeling can be fine-tuned 😊
- Accelerated ML will be able to play with larger datasets
- Accelerated ML saves CPU cycles , saves time & effort,  
& licensing costs (IBM i offload)
- Auto ML can create high quality models, that only  
experienced data scientists can build**



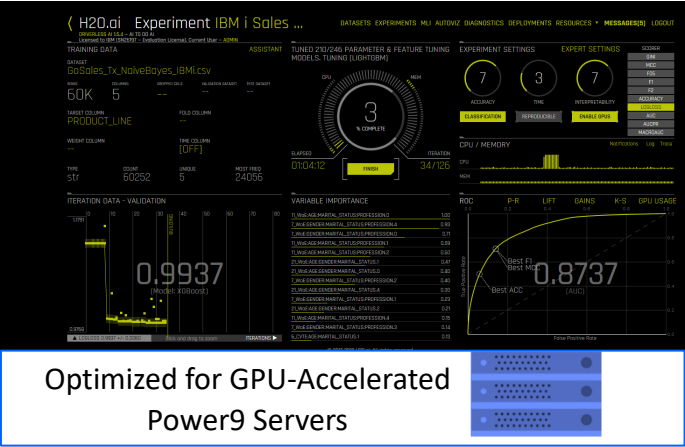
# Smarter IBM i apps made easy with Driverless AI

## CRM & Customer Churn scenario

# Summary

Import & Visualize Data  
Create & Test Models

Extract  
CRM data



Optimized for GPU-Accelerated  
Power9 Servers

Model Export



Recommendation Engine  
Scoring Pipeline

Augment IBM i Applications  
Model Inference

Monitor Models  
& Iterate



Business Database

```
[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"]}
```

Business Application

# Smarter IBM i apps made easy with Driverless AI

## CRM & Customer Churn scenario

Customer Relationship Management - IBM i + Machine Learning Driverless AI MOJO Scoring Pipeline

### Customers Dashboard

| CUSTOMERID | GENDER | STREAMI...         | CONTRACT        | PAPERLE... | PAYMENT...          | MONTHL... | TOTALCH... | CHURN? |
|------------|--------|--------------------|-----------------|------------|---------------------|-----------|------------|--------|
| 7590-VHVEG | Female | No                 | Month-to-mon... | Yes        | Electronic check    | 29.85     | 29.85      |        |
| 5575-GNVDE | Male   | No                 | One year        | No         | Mailed check        | 56.95     | 1889.50    |        |
| 3668-QPYBK | Male   | No                 | Month-to-mon... | Yes        | Mailed check        | 53.85     | 108.15     |        |
| 7795-CFOCW | Male   | No                 | One year        | No         | Bank transfer (...) | 42.30     | 1840.75    |        |
| 9237-HQITU | Female | No                 | Month-to-mon... | Yes        | Electronic check    | 70.70     | 151.65     |        |
| 9305-CDSKC | Female | Yes                | Month-to-mon... | Yes        | Electronic check    | 99.65     | 820.50     |        |
| 1452-KIOVK | Male   | Yes                | Month-to-mon... | Yes        | Credit card (au...  | 89.10     | 1949.40    |        |
| 6713-OKOMC | Female | No                 | Month-to-mon... | No         | Mailed check        | 29.75     | 301.90     |        |
| 7892-POOKP | Female | Yes                | Month-to-mon... | Yes        | Electronic check    | 104.80    | 3046.05    |        |
| 6388-TABGU | Male   | No                 | One year        | No         | Bank transfer (...) | 56.15     | 3487.95    |        |
| 9763-GRSKD | Male   | No                 | Month-to-mon... | Yes        | Mailed check        | 49.95     | 587.45     |        |
| 7469-LKBCI | Male   | No internet ser... | Two year        | No         | Credit card (au...  | 18.95     | 326.80     |        |
| 8091-TTVAX | Male   | Yes                | One year        | No         | Credit card (au...  | 100.35    | 5681.10    |        |
| 0280-XJGEX | Male   | Yes                | Month-to-mon... | Yes        | Bank transfer (...) | 103.70    | 5036.30    |        |
| 5129-JLPIS | Male   | Yes                | Month-to-mon... | Yes        | Electronic check    | 105.50    | 2686.05    |        |
| 3655-SNQYZ | Female | Yes                | Two year        | No         | Credit card (au...  | 113.25    | 7895.15    |        |
| 8191-XWSZG | Female | No internet ser... | One year        | No         | Mailed check        | 20.65     | 1022.95    |        |
| 9959-WOFKT | Male   | Yes                | Two year        | No         | Bank transfer (...) | 106.70    | 7382.25    |        |
| 4190-MFLUW | Female | No                 | Month-to-mon... | No         | Credit card (au...  | 55.20     | 528.35     |        |
| 4183-MYFRB | Female | No                 | Month-to-mon... | Yes        | Electronic check    | 90.05     | 1862.90    |        |
| 8779-QRDMV | Male   | No                 | Month-to-mon... | Yes        | Electronic check    | 39.65     | 39.65      |        |
| 1680-VDCWW | Male   | No internet ser... | One year        | No         | Bank transfer (...) | 19.80     | 202.25     |        |

RESET

SCORING

REFRESH



Initial CRM :

No Customer Churn risk estimate per customer

# Smarter IBM i apps made easy with Driverless AI

## CRM & Customer Churn scenario

Customer Relationship Management - IBM i + Machine Learning Driverless AI MOJO Scoring Pipeline

### Customers Dashboard

| CUSTOM...  | GENDER | STREAMI...         | CONTRACT        | PAPERLE... | PAYMENT...          | MONTHL... | TOTALCH... | CHURN?     |
|------------|--------|--------------------|-----------------|------------|---------------------|-----------|------------|------------|
| 7590-VHVEG | Female | No                 | Month-to-mon... | Yes        | Electronic check    | 29.85     | 29.85      | Scoring... |
| 5575-GNVDE | Male   | No                 | One year        | No         | Mailed check        | 56.95     | 1889.50    | Scoring... |
| 3668-QPYBK | Male   | No                 | Month-to-mon... | Yes        | Mailed check        | 53.85     | 108.15     | Scoring... |
| 7795-CFOCW | Male   | No                 | One year        | No         | Bank transfer (...) | 42.30     | 1840.75    | Scoring... |
| 9237-HQITU | Female | No                 | Month-to-mon... | Yes        | Electronic check    | 70.70     | 151.65     | Scoring... |
| 9305-CDSKC | Female | Yes                | Month-to-mon... | Yes        | Electronic check    | 99.65     | 820.50     | Scoring... |
| 1452-KIOVK | Male   | Yes                | Month-to-mon... | Yes        | Credit card (au...  | 89.10     | 1949.40    | Scoring... |
| 6713-OKOMC | Female | No                 | Month-to-mon... | No         | Mailed check        | 29.75     | 301.90     | Scoring... |
| 7892-POOKP | Female | Yes                | Month-to-mon... | Yes        | Electronic check    | 104.80    | 3046.05    | Scoring... |
| 6388-TABGU | Male   | No                 | One year        | No         | Bank transfer (...) | 56.15     | 3487.95    | Scoring... |
| 9763-GRSKD | Male   | No                 | Month-to-mon... | Yes        | Mailed check        | 49.95     | 587.45     | Scoring... |
| 7469-LKBCI | Male   | No internet ser... | Two year        | No         | Credit card (au...  | 18.95     | 326.80     | Scoring... |
| 8091-TTVAX | Male   | Yes                | One year        | No         | Credit card (au...  | 100.35    | 5681.10    | Scoring... |
| 0280-XJGEX | Male   | Yes                | Month-to-mon... | Yes        | Bank transfer (...) | 103.70    | 5036.30    | Scoring... |
| 5129-JLPIS | Male   | Yes                | Month-to-mon... | Yes        | Electronic check    | 105.50    | 2686.05    | Scoring... |
| 3655-SNQYZ | Female | Yes                | Two year        | No         | Credit card (au...  | 113.25    | 7895.15    | Scoring... |
| 8191-XWSZG | Female | No internet ser... | One year        | No         | Mailed check        | 20.65     | 1022.95    | Scoring... |

RESET

SCORING

Invoke H2O DAI Scoring Pipeline

REFRESH

Real time Scoring  
(Db2 CRM database)

# Smarter IBM i apps made easy with Driverless AI

## CRM & Customer Churn scenario

Customer Relationship Management - IBM i + Machine Learning Driverless AI MOJO Scoring Pipeline

### Customers Dashboard

| CUSTOM...  | GENDER | STREAMI...         | CONTRACT        | PAPERLE... | PAYMENT...          | MONTHL... | TOTALCH... | CHURN? |
|------------|--------|--------------------|-----------------|------------|---------------------|-----------|------------|--------|
| 7590-VHVEG | Female | No                 | Month-to-mon... | Yes        | Electronic check    | 29.85     | 29.85      | 0.52   |
| 5575-GNVDE | Male   | No                 | One year        | No         | Mailed check        | 56.95     | 1889.50    | 0.15   |
| 3668-QPYBK | Male   | No                 | Month-to-mon... | Yes        | Mailed check        | 53.85     | 108.15     | 0.35   |
| 7795-CFOCW | Male   | No                 | One year        | No         | Bank transfer (...) | 42.30     | 1840.75    | 0.12   |
| 9237-HQITU | Female | No                 | Month-to-mon... | Yes        | Electronic check    | 70.70     | 151.65     | 0.46   |
| 9305-CDSKC | Female | Yes                | Month-to-mon... | Yes        | Electronic check    | 99.65     | 820.50     | 0.48   |
| 1452-KIOVK | Male   | Yes                | Month-to-mon... | Yes        | Credit card (au...  | 89.10     | 1949.40    | 0.27   |
| 6713-OKOMC | Female | No                 | Month-to-mon... | No         | Mailed check        | 29.75     | 301.90     | 0.29   |
| 7892-POOKP | Female | Yes                | Month-to-mon... | Yes        | Electronic check    | 104.80    | 3046.05    | 0.37   |
| 6388-TABGU | Male   | No                 | One year        | No         | Bank transfer (...) | 56.15     | 3487.95    | 0.13   |
| 9763-GRSKD | Male   | No                 | Month-to-mon... | Yes        | Mailed check        | 49.95     | 587.45     | 0.25   |
| 7469-LKBCI | Male   | No internet ser... | Two year        | No         | Credit card (au...  | 18.95     | 326.80     | 0.17   |
| 8091-TTVAX | Male   | Yes                | One year        | No         | Credit card (au...  | 100.35    | 5681.10    | 0.20   |
| 0280-XJGEX | Male   | Yes                | Month-to-mon... | Yes        | Bank transfer (...) | 103.70    | 5036.30    | 0.31   |
| 5129-JLPIS | Male   | Yes                | Month-to-mon... | Yes        | Electronic check    | 105.50    | 2686.05    | 0.33   |
| 3655-SNQYZ | Female | Yes                | Two year        | No         | Credit card (au...  | 113.25    | 7895.15    | 0.21   |
| 8191-XWSZG | Female | No internet ser... | One year        | No         | Mailed check        | 20.65     | 1022.95    | 0.15   |
| 9959-WOFKT | Male   | Yes                | Two year        | No         | Bank transfer (...) | 106.70    | 7382.25    | 0.20   |
| 4190-MFLUW | Female | No                 | Month-to-mon... | No         | Credit card (au...  | 55.20     | 528.35     | 0.28   |
| 4183-MYFRB | Female | No                 | Month-to-mon... | Yes        | Electronic check    | 90.05     | 1862.90    | 0.29   |
| 8779-QRDMV | Male   | No                 | Month-to-mon... | Yes        | Electronic check    | 39.65     | 39.65      | 0.56   |
| 1680-VDCWW | Male   | No internet ser... | One year        | No         | Bank transfer (...) | 19.80     | 202.25     | 0.17   |

RESET

SCORING

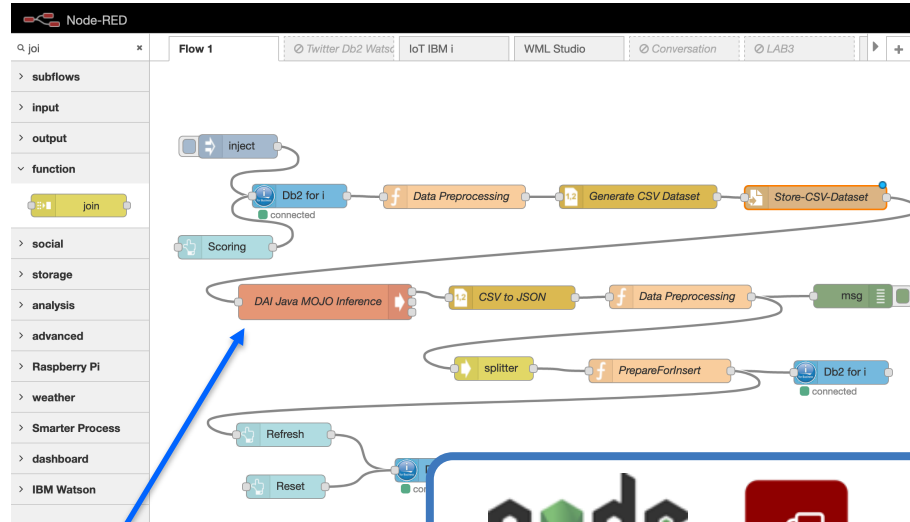
REFRESH

Enriched CRM Data  
With Churn Prediction

# Smarter IBM i apps made easy with Driverless AI

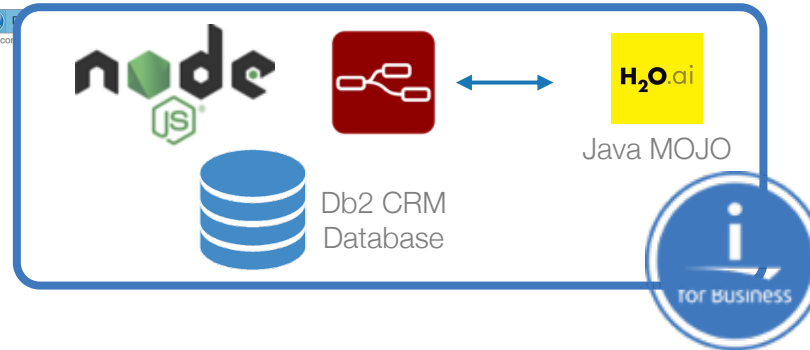
## CRM & Customer Churn scenario

Fig. Node-RED Prototyping - CRM dashboard



Behind the scenes:  
CRM application prototype  
integrated with a real time  
scoring Pipeline (Java MOJO)

Java MOJO Scoring Pipeline  
Invoked from the Node.js app



# Demonstration

*Accelerated Machine Learning with H2O.ai Driverless AI*



***Additional Content:***  
***Driverless AI Data Connectors & JDBC***  
***(Data Procurement / Preparation phase)***

<https://www.ibm.com/systems/clientcenterdemonstrations/faces/dcDemoView.jsp?demoId=3282>

# Smarter IBM i apps made easy with Driverless AI

## CRM & Customer Churn scenario

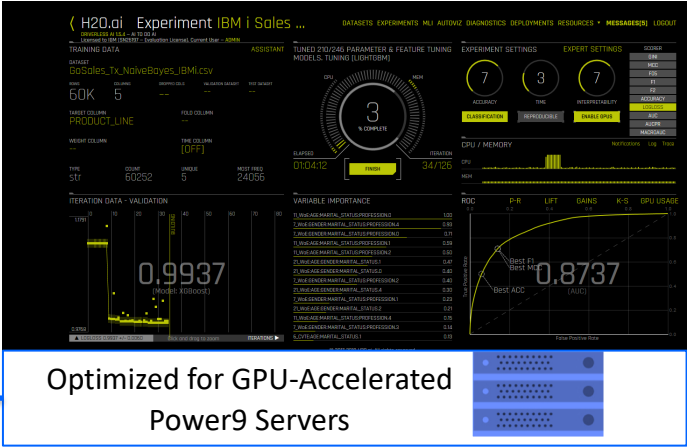
Summary  
with JDBC  
Connection

Import & Visualize Data  
Create & Test Models

Extract  
CRM data

Direct Db2  
Connection

JDBC



Model Export



Recommendation Engine  
Scoring Pipeline

Augment IBM i Applications  
Model Inference

Monitor Models  
& Iterate



Business Database

```
[14:57:33][api-appserver.power.com][/home/BENOIT/DAI]# uname -a
OS400 sp4i-appserver 2 7 00100006C37T
[14:57:37][api-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"]}
```

Business Application

# DAI Data Connectors: JDBC Connection to DB2 for i

## How to configure a Db2 Data Connector:

1. Download a Db2 JDBC Driver first
  - Ex: Db2 for i type 4 jdbc driver jt400.jar : <https://sourceforge.net/projects/jt400/>
  - Db2 LUW : get the appropriate db2jcc4.jar
2. Get the original Driverless AI Config file aka /etc/dai/config.toml  
Many ways to do that, either from the image or a running container (docker cp etc.)
  - K8s example : `kubectl cp -c dai-gpu-mop dai-demo-mop-6b75f6b5b9-kddd:/etc/dai .`
3. Edit this config.toml as shown below (refer to the Online doc for further information)
4. Inject config.toml, and the driver jar file(s) , to be mounted in your DAI POD/Docker container.
  - Inject it in your DAI image or better, use a K8s ConfigMap / persistent volume to persist it
5. Restart your DAI container to refresh the config.

/etc/dai/config.toml

```
jdbc: JDBC Connector, remember to configure JDBC below. (jdbc_app_configs)
enabled_file_systems = "upload, file, hdfs, s3, jdbc "
```

```
jdbc_app_configs = '{"db2fori": {"url": "jdbc:as400://<IBMi-IP>:", "jarpath":
"/etc/dai/jt400.jar", "classpath": "com.ibm.as400.access.AS400JDBCdriver"}}'
```

```
Note: here, the jtopen driver jtopen.jar is mounted in the DAI Container, in /etc/dai
```

# DAI Data Connectors: JDBC Connection to DB2 for i

< H2O.ai

DRIVERLESS AI 1.7.1 – AI TO DO AI  
Licensed to IBM (SN37526 – Evaluation License). Current User – ADMIN

PROJECTS DATASETS AUTOVIZ EXPERIMENTS DIAGNOSTICS MLI DEPLOYMENTS RESOURCES ▼ MESSAGES[4] LOGOUT

Datasets

| <input type="checkbox"/> | Name ↕                               | Path ↕                            | Size ↕ | Rows ↕ | Columns ↕ |
|--------------------------|--------------------------------------|-----------------------------------|--------|--------|-----------|
| <input type="checkbox"/> | TEST                                 | ...90/TEST.1568642468.7804182.bin | 963KB  | 7K     | 21        |
| <input type="checkbox"/> | Churn.csv                            | ...urn.csv.1566400172.8270016.bin | 1MB    | 7K     | 21        |
| <input type="checkbox"/> | UCI_Credit_Card.csv                  | ...ard.csv.1562228258.5946171.bin | 4MB    | 30K    | 25        |
| <input type="checkbox"/> | GoSales_Tx_NaiveBayes.csv            | ...ayes.csv.1558533238.553315.bin | 3MB    | 60K    | 5         |
| <input type="checkbox"/> | WA_Fn-UseC_-Telco-Customer-Churn.csv | ...urn.csv.1558036429.8231266.bin | 1MB    | 7K     | 21        |

+ ADD DATASET (OR DRAG & DROP)

FILE SYSTEM  
Status ↕

UPLOAD FILE  
[Click for Actions]

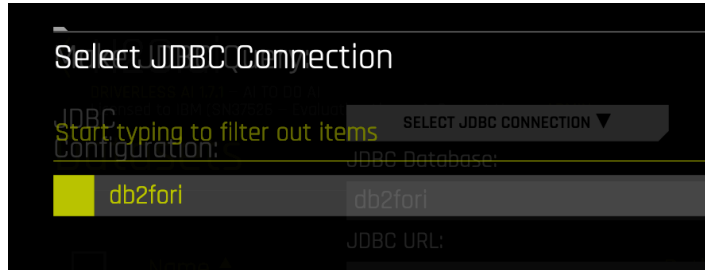
AMAZON S3  
[Click for Actions]

HADOOP FILE SYSTEM  
[Click for Actions]

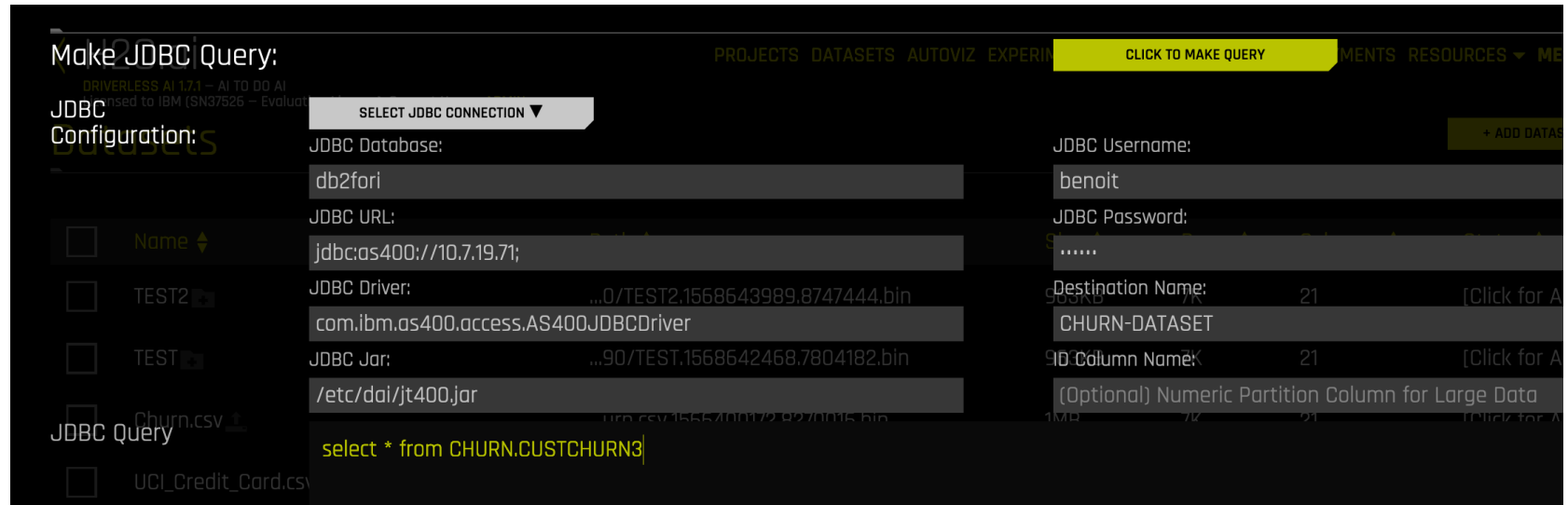
JDBC  
[Click for Actions]

Data Connectors - JDBC

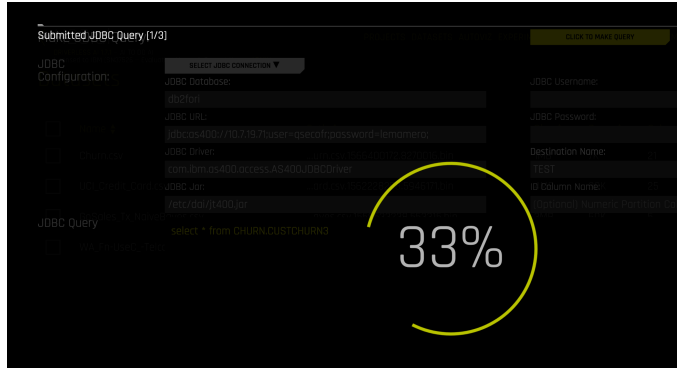
# DAI Data Connectors: JDBC Connection to DB2 for i



Configure your Data source:  
Db2 Connection & Query with credentials  
to your Db2 database.



# DAI Data Connectors: JDBC Connection to DB2 for i



Explore your dataset populated with fresh data from your Db2 for i database.

H2O.ai  
DRIVERLESS AI 1.7.1 – AI TO DO AI  
Licensed to IBM (SN37526 – Evaluation License), Current User – ADMIN

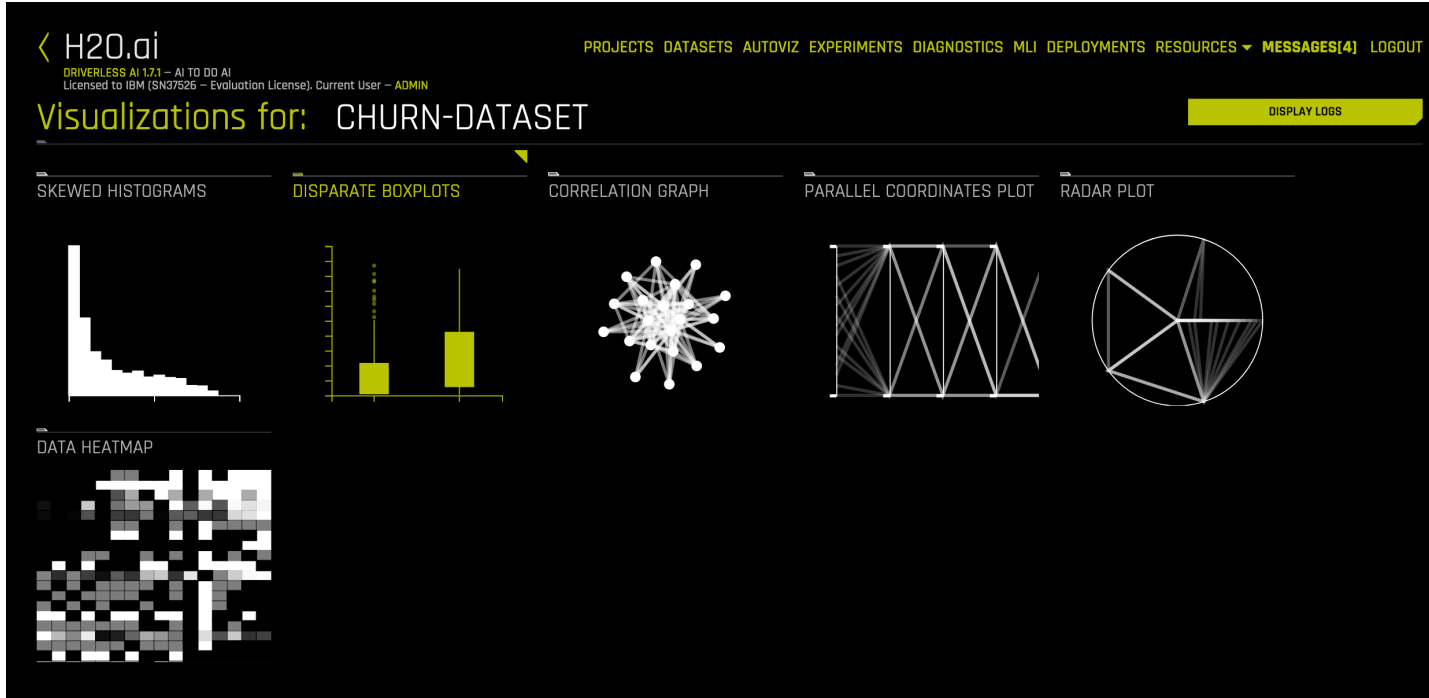
PROJECTS DATASETS AUTOVIZ EXPERIMENTS DIAGNOSTICS MLI DEPLOYMENTS RESOURCES ▼ MESSAGES[4] LOGOUT

+ ADD DATASET (OR DRAG & DROP)

| <input type="checkbox"/> | Name ↕                | Path ↕                            | Size ↕ | Rows ↕ | Columns ↕ | Status ↕            |
|--------------------------|-----------------------|-----------------------------------|--------|--------|-----------|---------------------|
| <input type="checkbox"/> | CHURN-DATASET 🚩       | ...N-DATASET.1568644115.70989.bin | 963KB  | 7K     | 21        | [Click for Actions] |
| <input type="checkbox"/> | TEST2 🚩               | ...0/TEST2.1568643989.8747444.bin | 963KB  | 7K     | 21        | [Click for Actions] |
| <input type="checkbox"/> | TEST 🚩                | ...90/TEST.1568642468.7804182.bin | 963KB  | 7K     | 21        | [Click for Actions] |
| <input type="checkbox"/> | Churn.csv 📄           | ...urn.csv.1566400172.8270016.bin | 1MB    | 7K     | 21        | [Click for Actions] |
| <input type="checkbox"/> | UCI_Credit_Card.csv 📄 | ...ard.csv.1562228258.5946171.bin | 4MB    | 30K    | 25        | [Click for Actions] |

# DAI Data Connectors: JDBC Connection to DB2 for i

Explore your dataset populated  
with fresh data from your Db2 for i database.



# Demonstration

*Accelerated Machine Learning with H2O.ai Driverless AI*

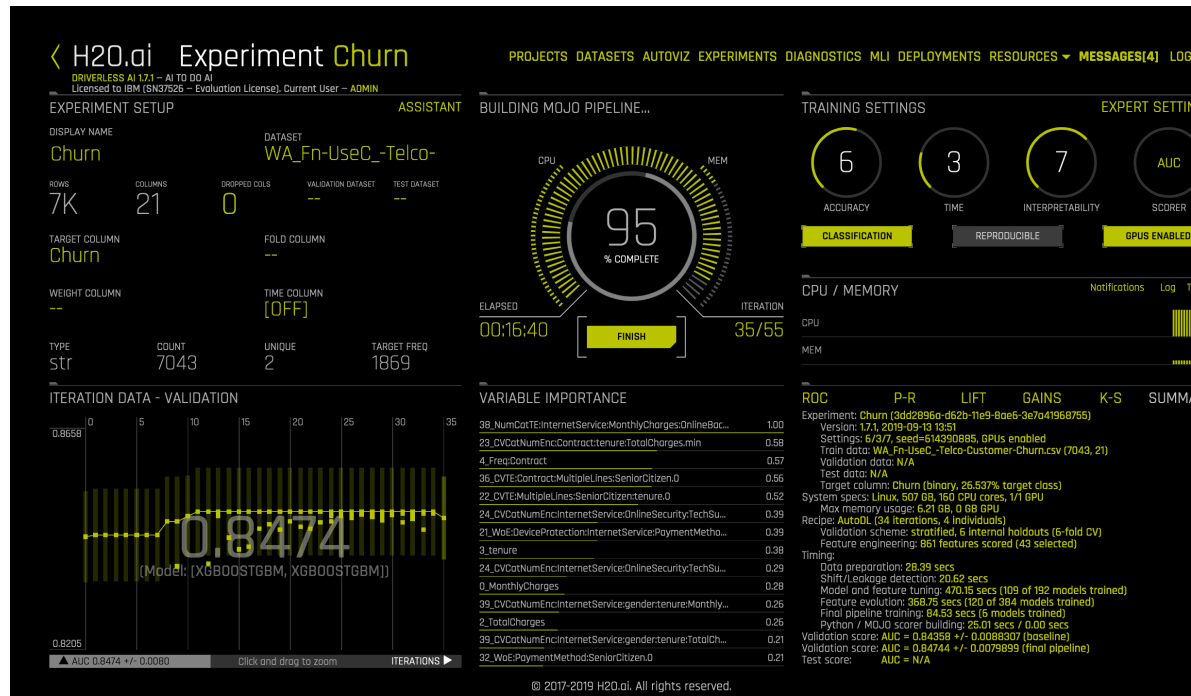
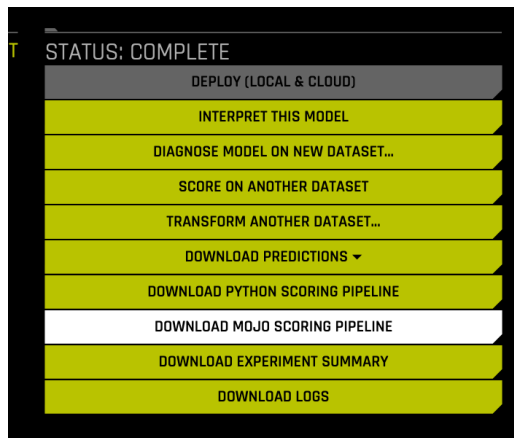


***Additional Content:***  
***Java Mojo Scoring Pipeline***  
***on IBM i/AIX/Linux***  
***(Model Inference, Deployment / Production phase)***

<https://www.ibm.com/systems/clientcenterdemonstrations/faces/dcDemoView.jsp?demoId=3282>

# Java Mojo Scoring Pipeline on IBM i/AIX/Linux

## Build and Export the MOJO Scoring Pipeline



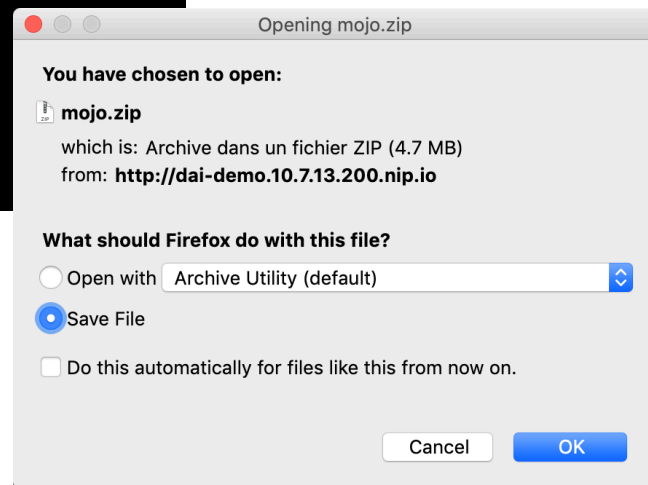
# Java Mojo Scoring Pipeline on IBM i/AIX/Linux

The screenshot displays the MOJO Scoring Pipeline interface. At the top, it says 'MOJO Scoring Pipeline instructions' and 'churn'. Below this, there's a navigation bar with 'PROJECTS', 'DATASETS', 'AUTOVIZ', 'EXPERIMENTS', 'DIAGNOSTICS', 'MLI', and 'DEPLOYMENTS'. The main area is titled 'EXPERIMENT SETUP' and shows the 'ASSISTANT' tab. The 'STATUS' is 'COMPLETE'. The 'TRAINING SETTINGS' section shows 'DISPLAY NAME: Churn', 'DATASET: WA\_F', 'LANGUAGE: JAVA', 'PYTHON', and 'R'. The 'TARGET COLUMN' is 'Churn'. The 'WEIGHT COLUMN' is '1'. The 'TYPE' is 'STF'. The 'COUNT' is '7043'. The 'UNDO' button is visible. The 'ITERATION DATA - VALIDATION' section shows a bar chart with values 0, 5, 10, 15, 20. A 'DOWNLOAD MOJO SCORING PIPELINE' button is at the bottom.

1. Unzip the downloaded mojo.zip  
2. cd mojo-pipeline  
3. bash ./run\_example.sh  
4. Follow instructions for MOJO Pipelines [from the documentation](#)

Download the generated  
Scoring Pipeline

*Customer Churn scoring pipeline*



# Java Mojo Scoring Pipeline on IBM i/AIX/Linux

```
[air-de-benoit:~ Benoît$ ssh benoit@10.7.19.71
[benoit@10.7.19.71's password:

***** IBM i 7.3 Montpellier Client Center *****

#####
#
#
#####
#
#
#####

***** Contact: benoit.marolleau@fr.ibm.com *****

[[10:22:18][BENOIT.ICC.LOCAL][~]# cd DAI/CustomerChurn/
[[10:22:32][BENOIT.ICC.LOCAL][~/DAI/CustomerChurn]# ls
MOJO scorer.zip scoring-pipeline
[[10:22:34][BENOIT.ICC.LOCAL][~/DAI/CustomerChurn]# cd MOJO/
[[10:22:38][BENOIT.ICC.LOCAL][~/DAI/CustomerChurn/MOJO]# ls
mojo.zip mojo2-runtime-2.1.4-all.jar
[[10:22:39][BENOIT.ICC.LOCAL][~/DAI/CustomerChurn/MOJO]# unzip mojo.zip
Archive: mojo.zip
 creating: mojo-pipeline/
 inflating: mojo-pipeline/README.txt
 inflating: mojo-pipeline/mojo2-runtime-javadoc.jar
 inflating: mojo-pipeline/mojo2-runtime.jar
 inflating: mojo-pipeline/run_example.sh
 inflating: mojo-pipeline/pipeline.mojo
 inflating: mojo-pipeline/example.csv
[[10:22:45][BENOIT.ICC.LOCAL][~/DAI/CustomerChurn/MOJO]#
```

Unzip the Scorer.zip file on the inference System.

Change the JVM version from 32 to 64 bit if necessary (necessary if max heap size  $\geq$  2.8GB) .

Heap size needed depends on the model size. Default 5GB

Documentation: *Typically, a good estimation for the amount of required memory is 12 times the size of the pipeline.mojo file*

```
[[10:29:39][BENOIT.ICC.LOCAL][~/DAI/CustomerChurn/MOJO/mojo-pipeline]# java -version
java version "1.8.0_201"
Java(TM) SE Runtime Environment (build 8.0.5.30 - pap3280sr5fp30-20190207_01(SR5 FP30))
IBM J9 VM (build 2.9, JRE 1.8.0 OS/400 ppc-32-Bit 20190124_408237 (JIT enabled, AOT enabled)
OpenJ9 - 9c77d86
OMR - dad8ba7
IBM - e2996d1)
JCL - 20190207_01 based on Oracle jdk8u201-b09
[[10:29:42][BENOIT.ICC.LOCAL][~/DAI/CustomerChurn/MOJO/mojo-pipeline]# export JAVA_HOME=/QOpenSys/QIBM/ProdData/JavaVM/jdk80/64bit
[[10:29:50][BENOIT.ICC.LOCAL][~/DAI/CustomerChurn/MOJO/mojo-pipeline]# java -version
java version "1.8.0_201"
Java(TM) SE Runtime Environment (build 8.0.5.30 - pap6480sr5fp30-20190207_01(SR5 FP30))
IBM J9 VM (build 2.9, JRE 1.8.0 OS/400 ppc64-64-Bit Compressed References 20190124_408237 (JIT enabled, AOT enabled)
OpenJ9 - 9c77d86
OMR - dad8ba7
IBM - e2996d1)
JCL - 20190207_01 based on Oracle jdk8u201-b09
```

# Java Mojo Scoring Pipeline on IBM i/AIX/Linux

Real time prediction on IBM i !! (AIX/Linux)  
with the java MOJO Scoring Pipeline built by Driverless AI

```
[11:21:14][BENOIT.ICC.LOCAL][~/DAI/CustomChurn/MOJO/mojo-pipeline]# uname -a
OS400 BENOIT 3 7 00100002BABV Os
[11:21:16][BENOIT.ICC.LOCAL][~/DAI/CustomChurn/MOJO/mojo-pipeline]# ./run_example.sh
+ MOJO_FILE=pipeline.mojo
+ CSV_FILE=example.csv
+ LICENSE_FILE=
+ CMD_LINE='java -Xmx5g -Dai.h2o.mojos.runtime.license.file=-cp mojo2-runtime-2.1.4-all.jar ai.h2o.mojos.ExecuteMojo'
+ cat
=====
Running MOJO2 example
=====

MOJO file : pipeline.mojo
Input file : example.csv

Command line : java -Xmx5g -Dai.h2o.mojos.runtime.license.file=-cp mojo2-runtime-2.1.4-all.jar ai.h2o.mojos.ExecuteMojo pipeline.mojo example.csv
+ java -Xmx5g -Dai.h2o.mojos.runtime.license.file=-cp mojo2-runtime-2.1.4-all.jar ai.h2o.mojos.ExecuteMojo pipeline.mojo example.csv
Mojo load time: 2.291 sec
Churn.No,Churn.Yes
0.8248027488589287,0.1751972511410713
0.8103899558385214,0.18961004416147867
0.9204178228974342,0.07958217710256577
0.691175510485967,0.30882448951403296
0.830219974120458,0.16978002587954202
0.9333716105255816,0.06662838947441843
0.8994874540302489,0.10051254596975114
0.7456717077228758,0.25432829227712417
0.6289845075872209,0.37101549241277904
0.3026417427592808,0.6973582572407192
Time per row: 4.100 msec (total time: 41.000 msec)
```

Real time Customer Churn predictions on IBM i

# Get Started Today : Contact us!

**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](mailto:a2roy@fr.ibm.com) / [benoit.marolleau@fr.ibm.com](mailto:benoit.marolleau@fr.ibm.com)

**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/>



# Montpellier Cognitive Systems Lab

Manager : Philippe Chonavel

Coordinator : Alain Roy

## Team PowerAI ATS Europe

Jean-Armand Broyelle

Regis Cely

Sebastien Chabrolles

Sylvain Delabarre

Maxime Deloche

Benoit Marolleau

## Team CAPI/SNAP (FPGA)

Bruno Mesnet

Alexandre Castelane

Fabrice Moyen

## Team HPC

Ludovic Enault

Pascal Vezolle

Montpellier Client Center (France)

## Our Offerings

Technical Consultancy & Assistance

Co-Creation Lab Workshops

Hands-On Technical Enablement

Demonstration/PoT

PoC/Benchmark

# Reference Architecture for AI Infrastructure

« Single Tenant »

AC922 – Internal disks



**PowerAI**

Cuda Drivers – OSS Frameworks

Red Hat Enterprise Linux (RHEL) or Ubuntu

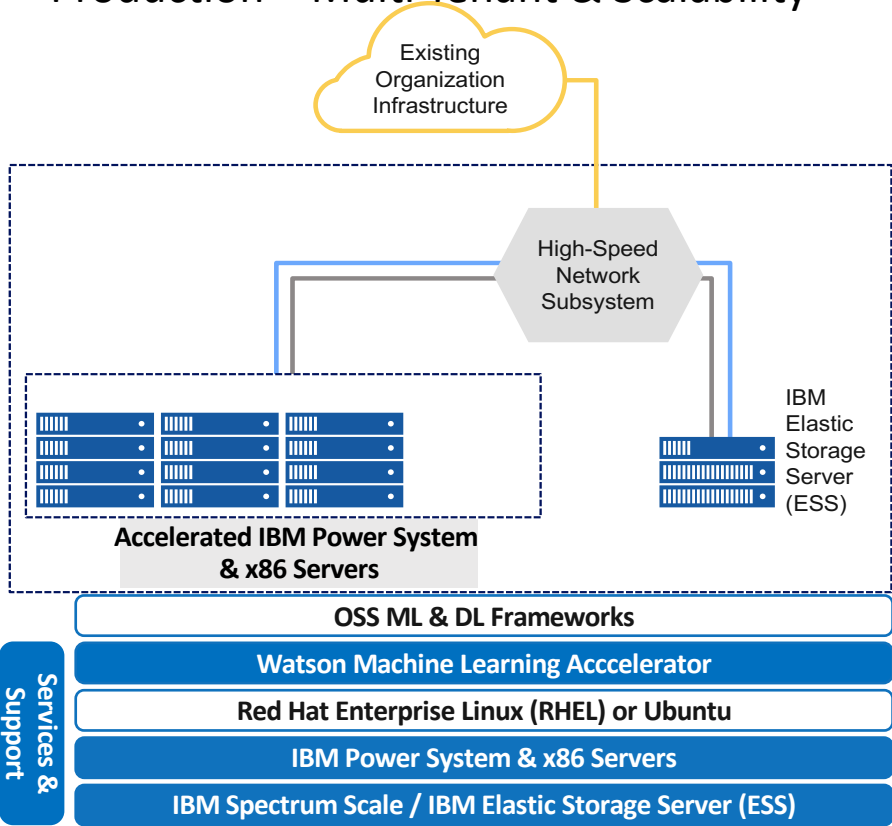


Internal SAS drives & NVM's

POWER Servers with GPU's

InfiniBand EDR P2P connection

Production – Multi Tenant & Scalability



|              |                         | 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                                                                  | Yes (AWS...)                                                                                                             |
|              | IBM Cloud Private       | Yes                                       | Yes                                                  | Yes                                           | Yes                                                                            | Yes                                                                                                                      |

Comparing AI Offerings on Power