IBM Systems
for Cognitive Solutions

IBM Machine Learning for z/OS

Khadija Souissi
IBM Client Center Boeblingen
Gartner identifies Machine Learning as the Top Trend in IT for 2017 and at the top of every CIO's strategy & budget

Source: Gartner

Machine learning segment of the cognitive computing market forecast to grow from $6 billion in 2016 to $52 billion in 2021 with a CAGR of 53.5% for 2016-2021

Published date: 05/02/2016 Source: Mindcommerce

Data scientists are the superheroes and unicorns of today's business. But data scientists are only human, and they are reaching the limits of productivity with current processes.

Published date: 02/11/2016 Source: Forrester
Machine learning is everywhere, influencing nearly everything we do…

Machine Learning

- Identifies patterns in historical data
- Builds behavioral models from patterns
- Makes recommendations
What is Machine Learning?

Machine Learning

- Supervised: Task-driven (Regression / Classification)
- Unsupervised: Data driven (Clustering)
- Reinforcement: Algorithm learns to react to an environment

Machine Learning

- Artificial Intelligence: Any technique which enables computers to mimic human behavior.
- Machine Learning: Subset of AI techniques which use statistical methods to enable machines to improve with experiences.
- Deep Learning: Subset of ML which make the computation of multi-layer neural networks feasible.

Deep Learning

Convolutional Neural Networks (CNN)
- Convolutions: Localized feature extraction
- Fully connected layers: Global feature integration

- Architecture:
  - L0 (Input)
  - L1: 512x512
  - L2: 256x256
  - L3: 128x128
  - L4: 64x64
  - L5: 32x32
  - F5: (Output)
Machine Learning Workflow: The Perception

Data ? Machine Learning Algorithm ? $
Machine Learning Workflow: The Reality

MANUAL INTERVENTION THROUGHOUT

Creating Examples

Choosing the Best Model

Predict

Models Lose Accuracy

Data

Data Prep

Machine Learning Algorithm

Model

Deploy

Scalable Deployment

Automating Data Science Work

MANUAL INTERVENTION THROUGHOUT

$
Create, deploy and manage behavioral models

Ingest Data → Extract Features → Train Model → Deploy Model → Make Predictions

Choose Best Model → Identify Model Degradation → Create Examples → Ensure Scalability

Human Intervention

Requires significant development, deployment, management and human intervention

The need for machine learning is surpassing the resources to optimize its use.
Challenges faced by data scientist ...

- Growing number of predictive models to be implemented
- Access various types of data
- More than 50 different algorithms: SVM, Neural Net, Decision Trees/Forests, Naïve Bayes, Regression, SMO, k-nearest Neighbor, Clustering, Rules, …
- Combinatorically explosive number of parameter choices per algorithm: kernel type, pruning strategy, number of trees in a forest, learning rate, …
- Wide variation in performance across different algorithm implementations (e.g., SPSS vs Python vs WEKA vs SPARK …)
- User-Defined algorithms
- Model Accuracy
IBM Machine Learning

Quick model development
Fast deployment
Continuous auditing & proactive notification
Easy management

Ingest Data → Extract Features → Train Model → Deploy Model → Make Predictions

Continuous feedback, Retraining

Data scientists can focus on the quality of the models and not the complexities of the process
Infuse continuous intelligence throughout the enterprise

**Freedom:** Choose the right language and ML framework and platform for your business.

**Productivity:** Make both experienced and novice data scientists more productive.

**Trust:** Confidently deploy insights knowing they are generated from the most current data and trends.
IBM Machine Learning: Extracted from Watson, delivered to our private cloud

- Cognitive computing
- Augmented intelligence
- Machine learning

IBM Watson → IBM PRIVATE CLOUD → IBM Machine Learning
IBM Systems: Analytics Infrastructure Roadmap

IBM z Systems

IBM Power Systems

Today

IBM Machine Learning for z/OS

Deep Learning Frameworks
- IBM Caffe
- Caffe
- NVCaffe

Python

Spark

R

Anaconda

IBM Machine Learning for Power*

IBM plans to deliver IBM Machine Learning on its IBM Power Systems
First Available for z/OS

- Where much of the world’s most valuable industry data runs
- Bring machine learning to your operational data
IBM Machine Learning (ML) for z/OS

- Consistent IBM offering
- On premise machine learning solution for z/OS
- Uses the same tools and technologies as the IBM Machine Learning (ML) Service
- Holistic approach, which offers data science experience (data prep, training and evaluation), model deployment, scoring, monitoring, and model retraining
- Can be used with and benefit from DB2 Analytics Accelerator
- Leverages z/OS Platform for Apache Spark (Apache Spark on z/OS)
IBM Machine Learning for z/OS – Faster Time-to-Value

<table>
<thead>
<tr>
<th>Business Value</th>
<th>Technology Value</th>
<th>Capability</th>
</tr>
</thead>
<tbody>
<tr>
<td>✓ Convert more opportunities into revenue</td>
<td>Create better models in less time</td>
<td>Cognitive Assistant for Data Scientists (CADS)</td>
</tr>
<tr>
<td>✓ Optimize resources, predictions and decisions</td>
<td>Simplify model creation</td>
<td>Hyper Parameter Optimization (HPO)</td>
</tr>
<tr>
<td>✓ Automatically identify and minimize risk</td>
<td>Improve models over time</td>
<td>DSX Pipeline User Interface</td>
</tr>
<tr>
<td>✓ Disrupt the competition and the disrupters</td>
<td>Easily integrate with existing tools and applications</td>
<td>Continuous Monitoring and Feedback Loop</td>
</tr>
<tr>
<td></td>
<td>Simplify model management</td>
<td>Modern RESTful APIs</td>
</tr>
</tbody>
</table>

- Rapidly optimize the algorithm that best fits the data and business scenario
- Provide optimal parameters for any given model
- Wizards make it easy for users to create and train a model
- Monitor model accuracy with feedback data and performance history
- Notification of model performance deterioration for more efficient retraining
- Ease collaboration across users (e.g., Data Scientists and App Developers)
- Easily manage thousands of models in an enterprise environment
IBM Machine Learning for z/OS – ML for the Enterprise

Model creation with Machine Learning UI & Integrated Notebook

Cognitive Assistant for Data Scientists (CADS) Hyper Parameter Optimization (HPO)

Multiple Model Deployment Options (Online scoring on z – REST API)

Model Management & Governance

Continuous Monitoring and Feedback
IBM Machine Learning for z/OS - Architecture and Deployment

- Linux
  - Machine Learning User Interface
  - Machine Learning Application

- z/OS
  - IBM z/OS Platform for Apache Spark
  - Data
IBM ML for z/OS – Current Architecture

Machine Learning User Interface (UI)

Application Cluster

- Metadata Service
- Deployment Service (Feedback, Evaluation, Monitoring)
- Jupyter Server
- Jupyter Kernel Gateway
- Apache Toree Proxy
- Ingestion Service
- Transformation Service
- Pipeline Service

Service Metadata, ML Models

- DB2
- zLDAP
- SDBM/LDBM
- DB2 JDBC

z/OS Spark Cluster

- Ingestion Lib
- Transformation Lib
- Feedback Lib
- Pipeline Lib
- Spark ML Model Lib
- CADS/HPO Lib
- DB2 JDBC

Data Source

- DB2
- IMS
- VSAM
- SMF

ODA

UI Metadata

- CouchDB

Authentication Service

Linux

Apache Toree

DB2

Scoring service

z/OS Spark in Local Mode

z/OS

Apache Toree Kernel

WebSphere Liberty

LDAP Lightweight Directory Access Protocol

NGINX

QUARTZ

CouchDB
IBM z/OS Platform for Apache Spark (Spark on z/OS)
Available since December 2015 via Open Source

Security:
- Integrate OLTP and Business Critical Data

Unique capability:
- Only found on Apache Spark on z/OS

Integrate:
- DB2 for z/OS, IMS, VSAM, PDSE, Syslog, SMF, ...
- Remote (non-z) data on distributed servers, Hadoop, Oracle, ...
- Defining security authorizations for instance using RACF
Why IBM machine Learning for z/OS

- **Data gravity**
  - Most data tends to already reside on z Systems; this “critical mass” is a key consideration for machine learning
  - Mirroring data to remote sites is not only costly but affects data currency with latency and multiple copies
  - For any data residing on distributed servers, IDAA can be instrumental and cost effective to bring data onto z Systems
- **Scoring Latency**
  - Predictive analytics requires prompt scoring results that can be imbedded in sub-second online transactions in real-time
  - Co-locating machine learning analytics with data on the same physical server enables immediate and accurate insights
- **Security**
  - Multiple copies of data can cause security and compliance concerns, especially for data required to move outside the firewall
Why IBM machine Learning for z/OS

- **Scalability and performance**
  - z Systems has the highest quality of services, with 99.999% availability
  - z Systems has extreme scalability with up to 141 processors (CP or IFL), 94 zIIPs, and up to 10TB memory in a single server

- **Automation and Life Cycle Management**
  - CADS (Cognitive Assistant for Data Scientists) and HPO (Hyper Parameter Optimization) accelerate ML model creation
  - ML's continuous learning loop to improves models over time
Machine learning can be applied to a variety of use cases – across problem types and industries

**Banking**
- **Card:** Using scoring to determine transaction risk based on user spending history.
- **Money Laundering:** Risk based on multiple account adversarial money wiring patterns.

**Retail**
- **In-Context Promotions:** Real-time scoring for personalized marketing.

**Government**
- **Compliance:** Scoring to detect non-compliant behavior and tax evasion.
- **Social Services:** Assessing likelihood that individuals will need multiple agency support to proactively engage various agencies to create best outcome and manage costs.
Continuous Intelligence in healthcare

Using machine learning to personalize prescription plans & lower payments

**THE PROBLEM**
Develop personalized prescription plan for diabetes patients based on their individual risk profile.

**SOLUTION**
Classify patients as low, moderate, high risk of developing diabetes based on blood sugar, blood pressure and cholesterol.

**Data and machine learning at work**
- Co-pay based on risk profile
- Real-time response at point of sale
- Ongoing tracking of risk profile changes
THANK YOU!
Khadija Souissi
IBM Client Center – Systems and Software – z ATS
IBM Germany Lab

Schoenaicher Str. 220
D-71032 Boeblingen

Phone +49 (0) 162 280 6856
+49 7031 16 5018

souissik@de.ibm.com
Backup
Data scientists love using notebooks

A notebook is a kind of living program that is interactive
Using Jupyter notebook – data selection

Load the data from DB2 for z/OS into dataframe and split for training, testing and evaluating

```python
In [2]:
import org.apache.spark.sql.SparkSession
import org.apache.spark.sql.functions_

//Load data from DB2 for z/OS using JDBC driver
val churnDataRaw = spark.read.format("jdbc").
    options(Map("driver" -> "com.ibm.db2.jcc.DB2Driver",
    "url" -> "jdbc:db2://192.168.1.1:430/LQCDB11",
    "user" -> "db2user01","password" -> "password",
    "dbtable" => "SA.CUST_GM").load()

val toDouble = udf {x: Int => x.toDouble}

val churnData = churnDataRaw.select("AGE", "ACTIVITY", "EDUCATION", "SEX", "STATE", "NEGTVTWEETS", "INCOME", "CHURN_LABEL")
churnData.show(5)
```

<table>
<thead>
<tr>
<th>AGE</th>
<th>ACTIVITY</th>
<th>EDUCATION</th>
<th>SEX</th>
<th>STATE</th>
<th>NEGTVTWEETS</th>
<th>INCOME</th>
<th>CHURN_LABEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>86</td>
<td>5</td>
<td>2</td>
<td>13857286.000</td>
<td>false</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Using Jupyter notebook – data split

```scala
val train = 80
val test = 10
val validate = 10

// Split the source dataset into training / testing / validation dataset
val splits = Sampling.trainTestSplit(df, train, test, validate)
val trainDF = splits._1
val testDF = splits._2
val validateDF = splits._3

println("schema of the source data:")
println(trainDF.schema)
trainDF.cache()
println("data preview:")
trainDF.show() //only showing top 3 rows
println(trainDF.count())
```

schema of the source data:
```text
StructType(StructField(label,DoubleType,true), StructField(GENDER,StringType,true), StructField(AGE,IntType,false), StructField(MARRITAL_STATUS,StringType,true), StructField(PROFESSION,StringType,true))
```

data preview:
```
+----------+--------+--------+----------------+--------+
| label    | GENDER | MARRITAL_STATUS | PROFESSION    |
|----------+--------+----------------+----------------+
| 0.0      | F      | Married        | Executive     |
| 0.0      | F      | Married        | Executive     |
| 0.0      | F      | Married        | Executive     |
|----------+--------+----------------+----------------+
```

only showing top 3 rows

17858
Using Jupyter notebook – model creation with CADS

```scala
// Feature definition
val genderIndexer = new StringIndexer(). setInputCol("SEX"). setOutputCol("gender_code")
val stateIndexer = new StringIndexer(). setInputCol("STATE"). setOutputCol("state_code")
val labelIndexer = new StringIndexer(). setInputCol("CHURN_LABEL"). setOutputCol("label")
val featuresAssembler = new VectorAssembler(). setInputCols(Array("AGE", "ACTIVITY", "EDUCATION", "MARRIED", "INCOME", "gender_code", "state_code"). setOutputCol("features")

// Select model automatically in candidate algorithms - Logistic Regression, SVM or Decision Tree?
val lr = new LogisticRegression(). setRegParam(0.05). setLabelCol("label"). setFeaturesCol("features")
val decisionTree = new DecisionTreeClassifier(). setMaxBins(50). setLabelCol("label"). setFeaturesCol("features")

// Cognitive Assistant for Data Scientists - predict model performance based on sampled data
val learners = List(LinearRegression(), decisionTree)
val cads = CADS(learner=new BinaryClassificationEvaluator(), setMetricName("areaUnderROC"). setNumFoldsParam(2). setNumSampleFoldParam(2). setEvaluator(learners). setKeepBestN(learners)(0). setTarget(Target(\"Pred\ prediction\", \"label\")).
val pipeline = new SparkPipeline().setStages(Array(labelIndexer, genderIndexer, stateIndexer, featuresAssembler, cads))
val model = pipeline.fit(trainingDF)
```
Using Jupyter notebook – model evaluation
Using a Visual Builder

- Unlocks the world of Data Science to none Data Scientists

- Allows Data Scientists to be more productive
Visual model builder - Data selection

<table>
<thead>
<tr>
<th>NAME</th>
<th>TYPE</th>
<th>SERVICE</th>
</tr>
</thead>
<tbody>
<tr>
<td>tent_data</td>
<td>Connection</td>
<td>DB2 for z/OS</td>
</tr>
<tr>
<td>Bank_Cux_Data</td>
<td>Connection</td>
<td>DB2 for z/OS</td>
</tr>
</tbody>
</table>
Visual model builder -- Model Deployment and Predict
Visual model builder – Mode Monitoring Dashboard
Monitoring Health by Evaluating Area under the ROC Curve

Monitoring the Area under the ROC Curve
Legal Disclaimer

- © IBM Corporation 2017. All Rights Reserved.
- The information contained in this publication is provided for informational purposes only. While efforts were made to verify the completeness and accuracy of the information contained in this publication, it is provided AS IS without warranty of any kind, express or implied. In addition, this information is based on IBM’s current product plans and strategy, which are subject to change by IBM without notice. IBM shall not be responsible for any damages arising out of the use of, or otherwise related to, this publication or any other materials. Nothing contained in this publication is intended to, nor shall have the effect of, creating any warranties or representations from IBM or its suppliers or licensors; or altering the terms and conditions of the applicable license agreement governing the use of IBM software.
- References in this presentation to IBM products, programs, or services do not imply that they will be available in all countries in which IBM operates. Product release dates and/or capabilities referenced in this presentation may change at any time at IBM’s sole discretion based on market opportunities or other factors, and are not intended to be a commitment to future product or feature availability in any way. Nothing contained in these materials is intended to, nor shall have the effect of, stating or implying that any activities undertaken by you will result in any specific sales, revenue growth or other results.
- If the text contains performance statistics or references to benchmarks, insert the following language; otherwise delete:
  - Performance is based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon many factors, including considerations such as the amount of multiprogramming in the user’s job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve results similar to those stated here.
- If the text includes any customer examples, please confirm we have prior written approval from such customer and insert the following language; otherwise delete:
  - All customer examples described are presented as illustrations of how those customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics may vary by customer.
- Please review text for proper trademark attribution of IBM products. At first use, each product name must be the full name and include appropriate trademark symbols (e.g., IBM Lotus® Sametime® Unyte®). Subsequent references can drop “IBM” but should include the proper branding (e.g., Lotus Sametime Gateway, or WebSphere Application Server). Please refer to procedures in the Marketing Guide for guidance on which trademarks require the ® or ™ symbol. Do not use abbreviations for IBM product names in your presentation. All product names must be used as adjectives rather than nouns. Please list all of the trademarks that you use in your presentation as follows; delete any not included in your presentation: IBM, the IBM logo, Lotus, Lotus Notes, Notes, Domino, Quickr, Sametime, WebSphere, UC2, PartnerWorld and Lotusphere are trademarks of International Business Machines Corporation in the United States, other countries, or both. Unyte is a trademark of WebDialogs, Inc. in the United States, other countries, or both. If you reference Adobe® in the text, please mark the first use and include the following; otherwise delete:
  - Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.
- If you reference Java™ in the text, please mark the first use and include the following; otherwise delete:
  - Java and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates. If you reference Microsoft® and/or Windows® in the text, please mark the first use and include the following, as applicable; otherwise delete:
  - Microsoft and Windows are trademarks of Microsoft Corporation in the United States, other countries, or both.
- If you reference Linux® in your presentation, please mark the first use and include the following; otherwise delete:
  - Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both. Other company, product, or service names may be trademarks or service marks of others.
- If the text/graphics include screenshots, no actual IBM employee names may be used (even your own). If your screenshots include fictitious company names (e.g., Renovations, Zeta Bank, Acme) please update and insert the following; otherwise delete:
  - All references to [insert fictitious company name] refer to a fictitious company and are used for illustration purposes only.
- If you reference Intel® and/or any of the following Intel products in the text, please mark the first use and include those trademarks and logos as appropriate; otherwise delete:
  - Intel, Intel Centrino, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.
- If you reference UNIX® in the text, please mark the first use and include the following; otherwise delete:
  - UNIX is a registered trademark of The Open Group in the United States and other countries.
- If the text/graphics include screenshots, no actual IBM employee names may be used (even your own). If your screenshots include fictitious company names (e.g., Renovations, Zeta Bank, Acme) please update and insert the following; otherwise delete:
  - All references to [insert fictitious company name] refer to a fictitious company and are used for illustration purposes only.
- If you reference Adobe® in the text, please mark the first use and include the following; otherwise delete:
  - Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.