RuleML+RR 2017

Machine Learning, Optimization and Rules: Time for Agility and Convergence

Eric Mazeran & Jean-François Puget – IBM
Notices and Disclaimers

Copyright © 2017 by International Business Machines Corporation (IBM). No part of this document may be reproduced or transmitted in any form without written permission from IBM.

U.S. Government Users Restricted Rights - Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM.

Information in these presentations (including information relating to products that have not yet been announced by IBM) has been reviewed for accuracy as of the date of initial publication and could include unintentional technical or typographical errors. IBM shall have no responsibility to update this information. THIS DOCUMENT IS DISTRIBUTED “AS IS” WITHOUT ANY WARRANTY, EITHER EXPRESS OR IMPLIED. IN NO EVENT SHALL IBM BE LIABLE FOR ANY DAMAGE ARISING FROM THE USE OF THIS INFORMATION, INCLUDING BUT NOT LIMITED TO, LOSS OF DATA, BUSINESS INTERRUPTION, LOSS OF PROFIT OR LOSS OF OPPORTUNITY. IBM products and services are warranted according to the terms and conditions of the agreements under which they are provided.

Any statements regarding IBM’s future direction, intent or product plans are subject to change or withdrawal without notice.

Performance data contained herein was generally obtained in a controlled, isolated environments. Customer examples are presented as illustrations of how those customers have used IBM products and the results they may have achieved. Actual performance, cost, savings or other results in other operating environments may vary.

References in this document to IBM products, programs, or services does not imply that IBM intends to make such products, programs or services available in all countries in which IBM operates or does business.

Workshops, sessions and associated materials may have been prepared by independent session speakers, and do not necessarily reflect the views of IBM. All materials and discussions are provided for informational purposes only, and are neither intended to, nor shall constitute legal or other guidance or advice to any individual participant or their specific situation.

It is the customer’s responsibility to insure its own compliance with legal requirements and to obtain advice of competent legal counsel as to the identification and interpretation of any relevant laws and regulatory requirements that may affect the customer’s business and any actions the customer may need to take to comply with such laws. IBM does not provide legal advice or represent or warrant that its services or products will ensure that the customer is in compliance with any law.
Part 1 : Time for Agility*

(*) No link with agile development methods in this presentation.
Observation #1: Decision Making Apps are (mostly) single technology-centric so far
The Global Wealth Investment Management division of Bank of America uses the power of rules to validate millions of customer account details daily with IBM Operational Decision Manager.

In Germany, 100 planners optimize the overall production of Continental interactively, collaboratively and concurrently thanks to IBM Decision Optimization Center.
Example: Taxi Dispatch (real customer example, simplified here)

The Taxi company waits a bit before assigning cars to customers...

Making decisions one at a time leads to a myopic effect

Gathering data and constraints to understand the ‘big picture’ creates the opportunity for better decisions
How might we try to solve the marketing campaign problem?

- For each campaign, a cost $C$ and an expected return $R$

What about:

- Sort campaigns according to decreasing return to cost ratio $R / C$
- Choose campaigns in this order until the 100 budget is exhausted
How might we try to solve the marketing campaign problem?

- For each campaign, a cost \( C \) and an expected return \( R \)

What about:

- Sort campaigns according to decreasing return to cost ratio \( R / C \)
- Choose campaigns in this order until the 100 budget is exhausted

<table>
<thead>
<tr>
<th>Revenue</th>
<th>Cost</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>39</td>
<td>20</td>
<td>1.95</td>
</tr>
<tr>
<td>17</td>
<td>9</td>
<td>1.89</td>
</tr>
<tr>
<td>30</td>
<td>16</td>
<td>1.88</td>
</tr>
<tr>
<td>26</td>
<td>14</td>
<td>1.86</td>
</tr>
<tr>
<td>22</td>
<td>12</td>
<td>1.83</td>
</tr>
<tr>
<td>20</td>
<td>11</td>
<td>1.82</td>
</tr>
<tr>
<td>36</td>
<td>20</td>
<td>1.80</td>
</tr>
<tr>
<td>34</td>
<td>19</td>
<td>1.79</td>
</tr>
</tbody>
</table>
How might we try to solve the marketing campaign problem?

- For each campaign, a cost $C$ and an expected return $R$

What about:

- Sort campaigns according to decreasing return to cost ratio $R / C$
- Choose campaigns in this order until the 100 budget is exhausted

<table>
<thead>
<tr>
<th>Revenue</th>
<th>Cost</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>39</td>
<td>20</td>
<td>1.95</td>
</tr>
<tr>
<td>17</td>
<td>9</td>
<td>1.89</td>
</tr>
<tr>
<td>30</td>
<td>16</td>
<td>1.88</td>
</tr>
<tr>
<td>26</td>
<td>14</td>
<td>1.86</td>
</tr>
<tr>
<td>22</td>
<td>12</td>
<td>1.83</td>
</tr>
<tr>
<td>20</td>
<td>11</td>
<td>1.82</td>
</tr>
<tr>
<td>36</td>
<td>20</td>
<td>1.80</td>
</tr>
<tr>
<td>34</td>
<td>19</td>
<td>1.79</td>
</tr>
</tbody>
</table>

Revenue | Cost | Profit
---------|------|-------
154      | 82   | 72    

Copyright © 2017 by International Business Machines Corporation (IBM).
How might we try to solve the marketing campaign problem?

- For each campaign, a cost $C$ and an expected return $R$

What about:

- Sort campaigns according to decreasing return to cost ratio $R / C$
- Choose campaigns in this order until the 100 budget is exhausted

This kind of technique embodying domain knowledge to build up a solution is called a **heuristic**.

Heuristics have a number of weaknesses

- They don't guarantee to find the best solution or a solution at all
- They require good domain knowledge to create
- They can be hard to adapt if the problem changes (new constraints, for example)
How might we try to solve the marketing campaign problem?

- For each campaign, a cost $C$ and an expected return $R$

There is a better solution, which you can find using an optimization engine.

<table>
<thead>
<tr>
<th>Revenue</th>
<th>Cost</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>39</td>
<td>20</td>
<td>1.95</td>
</tr>
<tr>
<td>17</td>
<td>9</td>
<td>1.89</td>
</tr>
<tr>
<td>30</td>
<td>16</td>
<td>1.88</td>
</tr>
<tr>
<td>26</td>
<td>14</td>
<td>1.86</td>
</tr>
<tr>
<td>22</td>
<td>12</td>
<td>1.83</td>
</tr>
<tr>
<td>20</td>
<td>11</td>
<td>1.82</td>
</tr>
<tr>
<td>36</td>
<td>20</td>
<td>1.80</td>
</tr>
<tr>
<td>34</td>
<td>19</td>
<td>1.79</td>
</tr>
</tbody>
</table>

Revenue | Cost | Profit
---|------|---
185 | 100 | 85

vs 72 previously

Copyright © 2017 by International Business Machines Corporation (IBM).
Use-case centric usually calls for techno. combinations

Use case:
• forecast sales,
• use the forecast to plan production,
• analyze how uncertainty in the forecast impacts the plan.

Tools:
• combine predictive analytics with (stochastic) optimization and uncertainty what-if analysis.
Observation #2: Machine Learning fuels Decision Making
Deep Learning challenges « Traditional » AI

ARTICLE

Mastering the game of Go with deep neural networks and tree search

We introduce a new approach to computer Go that uses ‘value networks’ to evaluate board positions and ‘policy networks’ to select moves. These deep neural networks are trained by a novel combination of supervised learning from human expert games, and reinforcement learning from games of self-play. We also introduce a new search algorithm that combines Monte Carlo simulation with value and policy networks. Using this search algorithm, our program AlphaGo achieved a 99.8% winning rate against other Go programs, and defeated the human European Go champion by 5 games to 0.”
Machine Learning fuels Decision Making

(Outside) World

Decide
- Optimization
- Rules
- ...

Sense
- Machine Learning
- Natural Language Processing
- ...

Interpret / Predict
- Machine Learning
- Rules
- Knowledge Mgt
- ...

Structured insights

Structured data

Actions

Unstructured data
Observation #3: Analytics Wave boosts Decision Making
Clean Data, Comprehensive Data, Big Data

- Wrong data used to make Rule & Optimization engines fail.
- Incomplete data used to delay adoption of Decision Making systems.
- Leveraging Big Data technology, Rule systems become more pervasive.

Decision Services in Apache Spark/Hadoop
The Analytics Picture

Prescriptive Analytics
*What should I do?*

Diagnostic Analytics
*Why did it happen?*

Predictive Analytics
*What will happen?*

Descriptive Analytics
*What has happened?*

Analytics Focus

- Past
- Present
- Future

Source: [http://ibm.co/1gJyfl3](http://ibm.co/1gJyfl3)
Observation #4: Optimization is key to Prescriptive Analytics
Decision Optimization finds a feasible set of decisions such that, once applied, business objectives will be optimally achieved.

- Solve combinatorial problems that cannot be solved efficiently otherwise.
- Create the best possible plans.
Typical Prescriptive Analytics « Decisions for Actions »

• Choose (the best options among a set of possibilities)
• Assign (the resources to the tasks)
• Schedule (the tasks)
• Dispatch (the resources at the appropriate locations)
• Plan (what will be processed, for what purpose, and when)
• Configure (a set of parts as the appropriate artefact)

Copyright © 2017 by International Business Machines Corporation (IBM).
Optimization Supplements ML or Rules with Holistic Reasoning

Optimization take into account the several interacting parts of a system as a whole (holistically) including the many different types of relationships between them.

Example: Put «Predictive Maintenance» into Action

Given the current and estimated operating conditions of a piece of equipment, with ML we can predict the likelihood of the failure of the piece at a given date.

Optimization is required to
- give the best course of actions for executing the maintenance tasks given the predictions
- propose the best tradeoffs to minimize disruption of operations given the prediction and maintenance options

Copyright © 2017 by International Business Machines Corporation (IBM).
1. Industry/business use cases require to be agile in combining several technology.


3. The Analytics wave boosts decision making combining descriptive, predictive and prescriptive capabilities.

4. Value & ROI are brought by prescriptions. Rules and ML pick-up the pieces and Optimization assembles the puzzle.
Machine Learning, Optimization and Rules

Part 2 : Time for Convergence
Time for Convergence

We are working on 2 tracks:

- Common Machine Learning and Optimization workflow & algorithms
- Optimization « as Rules » for ease of modelling
Track #1:
Common Machine Learning and Optimization workflow
Basic Machine Learning

Training

Labeled examples

Feature Engineering → Training → Model

Data Scientist Time

Operations Time

Deploy

Scoring

New data

Feature Engineering → Scoring → Predicted data

Copyright © 2017 by International Business Machines Corporation (IBM).
The Decision Optimization cycle

What used to take days or weeks can now be achieved almost instantly ...

... and that allows business users to execute multiple what-if scenarios

\[
\min c^T x \\
\text{s.t. } Ax \leq b \\
x \text{ integer}
\]
Machine Learning = Optimization?

Create a model → Model → Solve a problem → Plan

Optimization

Train a model → Model → Score an instance → Action

Copyright © 2017 by International Business Machines Corporation (IBM).
The Analytics Workflow

Similar workflow for:
- Machine Learning
- Optimization
- And more…
The Analytics Workflow

Data
Prepare Data → Build models → Select Models → Deploy

Model Dev

Bus Analyst → Bus Analyst → Data Scientist → Bus Analyst → Bus Analyst

App Dev

Data Engineer

LOB

Bus Analyst

App Dev

Sandbox

Deploy → Evaluate → Deploy → Use models → Log → Monitor

Production

Blue collars

Bus Analyst

LOB

Bus Analyst

App Dev

Feedback

Copyright © 2017 by International Business Machines Corporation (IBM).
Outcomes:

- **Short term**: Improve efficiency and robustness of classifiers (better classifier regulation & create supersparse classifiers).
- **Long term**: «learning under constraints» (force learning to accept boundaries; avoid biases & improve interpretability).
Track #2: Optimization « as Rules » for ease of modelling
CURRENT OPTIMIZATION REQUIRES

Experts…
CURRENT OPTIMIZATION REQUIRES EXPERTS

... in maths (Operations Research)…
CURRENT OPTIMIZATION REQUIRES EXPERTS IN MATHS

... to embed the appropriate Optimization model in the application

\[
\text{minimize} \quad \sum_i y_i, \sum_j x_j
\]

subject to
\[
\forall i \sum_i x_i = 1
\]
\[
\forall ij \; x_i \leq y_j
\]
\[
\forall j \sum_j y_j \leq C_j
\]
Let’s eliminate this bottleneck and bring easily the benefits of prescriptive analytics to many more businesses.
It generates the appropriate optimization model for your combinatorial problem, relying only on your data, domains knowledge, and your natural language descriptions.
She also inspects the distance traveled by each of his reps... and notices that the allocation is unfair to some.
Bridget wants to start by focusing on making travel fair for her team, so she removes the default goal to replace it with a travel-related goal.
EXAMPLE : SALES TERRITORY ASSIGNMENT

She directly selects a suggestion about minimizing the distance covered by salesrep.
EXAMPLE : SALES TERRITORY ASSIGNMENT

She also wants to limit the number of states assigned to John, and for that she uses natural language input to get new suggestions.
EXAMPLE: SALES TERRITORY ASSIGNMENT

She selects the correct rule among the new suggestions.
She then selects "Compute refined plan" to find a new assignment based on the new goal of minimizing average travel per sales rep.
Using a bar chart, she can compare the two scenarios. She sees how the "Distance" scenario has equalized travel.
Explicit Model-based prescriptive decision making

Business Data (Inputs)

Prescriptive Decision Model

Apply the decision model on the inputs to propose the best actions.
Explicit Model-based prescriptive decision making

Rules-based (aka Decision Management)

- Rule Set
  - Capture decisional procedures. Specify decision cases and alternatives.
  - Choose the appropriate decisions, applying business rules on inputs.

Predefined decision policies: knows how to solve the problem

Optimization-based (aka Decision Optimization)

- Optimization
  - Define Decision Variables / Unknowns
  - Define Objectives and Constraints
  - Finds, given the inputs, a feasible set of decisions such that, once applied, business objectives will be optimally achieved.

Description of solution properties: does not know how to solve the problem.

Business Data (Inputs)

Copyright © 2017 by International Business Machines Corporation (IBM).
Explicit Model-based prescriptive decision making

- **Rules-based** (aka Decision Management)
  - Business Data (Inputs)
  - Rule Set
  - Decision Rules independent of the Implementation Logic
  - Choose the appropriate decisions, applying business rules on inputs.
  - Predefined decision policies: knows how to solve the problem

- **Optimization-based** (aka Decision Optimization)
  - Business Data (Inputs)
  - Optimization
  - Description of solution properties: does not know how to solve the problem.
  - Finds, given the inputs, a feasible set of decisions such that, once applied, business objectives will be optimally achieved.
Thank You!