Success Stories for Applying Big Data Analytics and Machine Learning to Real-time Processing

Kai Wähner

#TIBCONOW
SAFE HARBOR DISCLOSURE

During the course of this presentation, TIBCO or its representatives may make forward-looking statements regarding future events, TIBCO’s future results or our future financial performance. Although we believe that the expectations reflected in the forward-looking statements contained in this presentation are reasonable, these expectations or any of the forward-looking statements could prove to be incorrect and actual results or financial performance could differ materially from those stated herein.

TIBCO could experience factors that could cause actual results or financial performance to differ materially from those contained in any forward-looking statement made in connection with this presentation. TIBCO does not undertake to update any forward-looking statements that may be made from time to time or on its behalf.
DISCLAIMER

This document (including, without limitation, any product roadmap or statement of direction data) illustrates the planned testing, release and availability dates for TIBCO products and services. This document is provided for informational purposes only and its contents are subject to change without notice. TIBCO makes no warranties, express or implied, in or relating to this document or any information in it, including, without limitation, that this document, or any information in it, is error-free or meets any conditions of merchantability or fitness for a particular purpose. This document may not be reproduced or transmitted in any form or by any means without our prior written permission.

The material provided is for informational purposes only, and should not be relied on in making a purchasing decision. The information is not a commitment, promise or legal obligation to deliver any material, code, or functionality. The development, release, and timing of any features or functionality described for our products remains at our sole discretion.

During the course of this presentation TIBCO or its representatives may make forward-looking statements regarding future events, TIBCO’s future results or our future financial performance. These statements are based on management’s current expectations. Although we believe that the expectations reflected in the forward-looking statements contained in this presentation are reasonable, these expectations or any of the forward-looking statements could prove to be incorrect and actual results or financial performance could differ materially from those stated herein. TIBCO does not undertake to update any forward-looking statement that may be made from time to time or on its behalf.

© Copyright 2000-2016 TIBCO Software Inc. All rights reserved. TIBCO Confidential & Proprietary Information.
The following information is proprietary information of TIBCO Software Inc. Use, duplication, transmission, or republication for any purpose without the prior written consent of TIBCO is expressly prohibited.
Key Takeaways

- Machine Learning creates analytic models to discover new insights, which humans cannot find.
- Event Processing uses analytic models - without redeveloping - to act on future events in real time.
- Machine Learning enables huge new opportunities for business.
Agenda

1. Introduction to Machine Learning
2. Complete Fast Data Architecture
3. Success Story: Customer Experience and Cross Selling
4. Success Story: Predictive Maintenance
5. Live Demo
Agenda

1. Introduction to Machine Learning
2. Complete Fast Data Architecture
3. Success Story: Customer Experience and Cross Selling
4. Success Story: Predictive Maintenance
5. Live Demo
Machine learning is a method of data analysis that automates analytical model building. Using algorithms that iteratively learn from data, machine learning allows computers to find hidden insights without being explicitly programmed where to look.

http://www.sas.com
Real World Examples of Machine Learning

Spam Detection

Search Results + Product Recommendation

Picture Detection (Friends, Locations, Products)

Machine Learning is already present in your daily life...

Now, every enterprise is beginning to leverage it!
Data Scientists work with **many** Tools

- TIBCO(R) Enterprise Runtime for R
- Spark
- H2O.ai
- TERR
- Many more ....

© Copyright 2000-2016 TIBCO Software Inc. All rights reserved. TIBCO Confidential & Proprietary Information.
Agenda

1. Introduction to Machine Learning
2. Complete Fast Data Architecture
3. Success Story: Customer Experience and Cross Selling
4. Success Story: Predictive Maintenance
5. Live Demo
This document (including, without limitation, any product roadmap or statement of direction data) illustrates the planned testing, release, and availability dates for TIBCO products and services. It is for informational purposes only and its contents are subject to change without notice.

© Copyright 2000-2016 TIBCO Software Inc. All rights reserved. TIBCO Confidential & Proprietary Information.
This document (including, without limitation, any product roadmap or statement of direction data) illustrates the planned testing, release and availability dates for TIBCO products and services. It is for informational purposes only and its contents are subject to change without notice.

© Copyright 2000-2016 TIBCO Software Inc. All rights reserved. TIBCO Confidential & Proprietary Information.
Agenda

1. Introduction to Machine Learning
2. Complete Fast Data Architecture
3. Success Story: Customer Experience and Cross Selling
4. Success Story: Predictive Maintenance
5. Live Demo
**Scenario: Context-Specific Cross-Selling**

**Goal:** Cross-Sell perishable resources and make customers happy.

**Question:** What to sell to which customer in what situation?

You received a SMS from HappyAir:

Want to make your day more comfortable?

Spend some Loyalty Points for

- Lounge Access (A)
- Free Drinks on Board (B)
- Business upgrade (C)

Send a text message with A, B or C to 112233

or call 123-5555
This document (including, without limitation, any product roadmap or statement of direction data) illustrates the planned testing, release and availability dates for TIBCO products and services. It is for informational purposes only and its contents are subject to change without notice.

© Copyright 2000-2016 TIBCO Software Inc. All rights reserved. TIBCO Confidential & Proprietary Information.

Fast Data Architecture for Cross-Selling

- **FLIGHT DATA**
- **LOUNGE DATA**
- **GEOLOCATION**
- **SOCIAL FEEDS**

**Data Storage**
- Oracle (later Hadoop)

**BusinessWorks**
- SOA
- CRM
- Loyalty System

**BusinessEvents**
- Aggregate
- Correlate
- Analytics
- Rules

**Live Datamart**
- Continuous query processing
- Alerts
- Manual action, escalation

**Operational Analytics**
- Live UI

**HISTORICAL ANALYSIS**
- Spotfire
  - SAS
  - R / TERR
  - Python

**Integration Bus**

**Action**

**Operations**

**Data Scientists**

**TIBCO Fast Data Platform**

#TIBCONOW
TIBCO Spotfire with R / TERR Integration

Let the Business User leverage Analytic Models (created by the Data Scientist) to find insights!

Example: Cross Selling to avoid Customer Churn (R / TERR script “under the hood”)

Select features for re-computing the analytic model
R / TERR in BusinessEvents

Out-of-the-Box Support for R / TERR
Integration of any analytic model into BusinessEvents

**Implementation Steps**

1. **Model Registration**
2. **Events (Features and Model Scoring Result)**
3. **Model Scoring**
4. **Application of Scoring Result**

---

This document (including, without limitation, any product roadmap or statement of direction data) illustrates the planned testing, release and availability dates for TIBCO products and services. It is for informational purposes only and its contents are subject to change without notice. © Copyright 2000-2016 TIBCO Software Inc. All rights reserved. TIBCO Confidential & Proprietary Information.
Step-by-Step Introduction of Big Data Analytics

Step 1: Configure rules for BusinessEvents manually via web browser based on Spotfire / TERR data discovery (monthly)

Step 2: Push new rules from analytic models computed in Hadoop to BusinessEvents automatically (daily)

Step 3: Deploy and execute analytic models directly within engine of BusinessEvents (real time)
Agenda

1. Introduction to Machine Learning
2. Complete Fast Data Architecture
3. Success Story: Customer Experience and Cross Selling
4. **Success Story: Predictive Maintenance**
5. Live Demo
**Goal:** Scrap parts as early as possible to reduce costs in a manufacturing process.

**Question:** When to scrap a part in Station 1 instead of sending it to Station 2?

<table>
<thead>
<tr>
<th>Station 1</th>
<th>Station 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost Before</td>
<td>7€</td>
</tr>
<tr>
<td>Scrap?</td>
<td>Scrap?</td>
</tr>
</tbody>
</table>

Total Cost: 29€ (or more)
Fast Data Architecture for Predictive Maintenance

StreamBase
- Aggregate
- Correlate
- Analytics
- Rules

Live Datamart
- Continuous query processing
- Alerts
- Manual action, escalation

Operational Analytics

CSV Batch
JSON Real Time
XML Real Time

Action

Flume
HDFS

HISTORICAL ANALYSIS
- R/TERR
- H2O
- PMML

Live UI

Spotfire

TIBCO Fast Data Platform

Operations

Data Scientists

Oracle RDBMS

Hadoop (Cloudera)

Avro Parquet...

Internal Data

HDFS

This document (including, without limitation, any product roadmap or statement of direction data) illustrates the planned testing, release and availability dates for TIBCO products and services. It is for informational purposes only and its contents are subject to change without notice. © Copyright 2000-2016 TIBCO Software Inc. All rights reserved. TIBCO Confidential & Proprietary Information.

#TIBCONOW
TIBCO Spotfire with H2O Integration

Data Discovery / Data Mining ("Are parts that repeat a station more likely scrap parts?")
### Advanced Analytics (“Scrap parts as early as possible!”)

<table>
<thead>
<tr>
<th><strong>Run an H2O Random Forest model</strong></th>
<th><strong>H2O Random Forest Results</strong></th>
<th><strong>H2O GLM results</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Classification model: detect parts as either faulty or non-faulty</td>
<td><strong>Savings:</strong> 97,492 euros (29%)</td>
<td><strong>Savings:</strong> 78,328 euros (24%)</td>
</tr>
</tbody>
</table>

- **Select predictor variables**
  - Sensor1,s
  - Sensor1,n
  - Sensor2,s
  - Sensor2,n
  - Sensor3,s
  - Sensor3,n
  - Sensor4,s
  - Sensor4,n

- **Select number of trees**

- **Run the model**
  - Run H2O Random Forest
  - Run H2O GLM Model

<table>
<thead>
<tr>
<th><strong>H2O Random Forest Predictor Importance</strong></th>
<th><strong>H2O GLM Predictor Importance</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensor1,s</td>
<td>Sensor1,n</td>
</tr>
<tr>
<td>Sensor2,s</td>
<td>Sensor2,n</td>
</tr>
<tr>
<td>Sensor3,s</td>
<td>Sensor3,n</td>
</tr>
<tr>
<td>Sensor4,s</td>
<td>Sensor4,n</td>
</tr>
</tbody>
</table>

© Copyright 2000-2016 TIBCO Software Inc. All rights reserved. TIBCO Confidential & Proprietary Information.
### TIBCO Spotfire with H2O Integration

#### Advanced Analytics ("Scrap parts as early as possible!")

<table>
<thead>
<tr>
<th>H2O Random Forest Results</th>
<th>H2O GLM results</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Confusion Matrix:</strong></td>
<td><strong>Confusion Matrix:</strong></td>
</tr>
<tr>
<td>TRUTH</td>
<td>TRUTH</td>
</tr>
<tr>
<td>non-faulty</td>
<td>non-faulty</td>
</tr>
<tr>
<td>126,943</td>
<td>122,041</td>
</tr>
<tr>
<td>faulty</td>
<td>faulty</td>
</tr>
<tr>
<td>9,062</td>
<td>13,964</td>
</tr>
<tr>
<td></td>
<td>7,193</td>
</tr>
<tr>
<td></td>
<td>21,982</td>
</tr>
<tr>
<td><strong>Metrics:</strong></td>
<td><strong>Metrics:</strong></td>
</tr>
<tr>
<td>Accuracy</td>
<td>Accuracy</td>
</tr>
<tr>
<td>(TP + TN) / (TP + TN + FN)</td>
<td>(TP + TN) / (TP + TN + FN)</td>
</tr>
<tr>
<td>89.9 %</td>
<td>89.4 %</td>
</tr>
<tr>
<td>Precision</td>
<td>Precision</td>
</tr>
<tr>
<td>TP / (TP + FP)</td>
<td>TP / (TP + FP)</td>
</tr>
<tr>
<td>66.4 %</td>
<td>61.2 %</td>
</tr>
<tr>
<td>Recall</td>
<td>Recall</td>
</tr>
<tr>
<td>TP / (TP + FN)</td>
<td>TP / (TP + FN)</td>
</tr>
<tr>
<td>71.4 %</td>
<td>87.5 %</td>
</tr>
<tr>
<td>Specificity</td>
<td>Specificity</td>
</tr>
<tr>
<td>TN / (TN + FP)</td>
<td>TN / (TN + FP)</td>
</tr>
<tr>
<td>93.3 %</td>
<td>89.7 %</td>
</tr>
<tr>
<td><strong>Timing:</strong></td>
<td><strong>Timing:</strong></td>
</tr>
<tr>
<td>Model Training Time:</td>
<td>Model Training Time:</td>
</tr>
<tr>
<td>0.1 mins</td>
<td>0.1 mins</td>
</tr>
<tr>
<td>Script Run Time:</td>
<td>Script Run Time:</td>
</tr>
<tr>
<td>0.5 mins</td>
<td>0.5 mins</td>
</tr>
<tr>
<td><strong>Savings:</strong></td>
<td><strong>Savings:</strong></td>
</tr>
<tr>
<td>97,095 euros (30 %)</td>
<td>76,306 euros (23 %)</td>
</tr>
</tbody>
</table>

This document (including, without limitation, any product roadmap or statement of direction data) illustrates the planned testing, release and availability dates for TIBCO products and services. It is for informational purposes only and its contents are subject to change without notice.

© Copyright 2000-2016 TIBCO Software Inc. All rights reserved. TIBCO Confidential & Proprietary Information.
TIBCO Live Datamart

Operational Intelligence (“Monitor the manufacturing process and change rules in real time!”)

Live Dartmart Desktop Client
TIBCO Live Datamart

Operational Intelligence ("Monitor the manufacturing process and change rules in real time!")
Agenda

1. Introduction to Machine Learning
2. Complete Fast Data Architecture
3. Success Story: Customer Experience and Cross Selling
4. Success Story: Predictive Maintenance
5. Live Demo
Live Demo

TIBCO Spotfire + StreamBase + Live Datamart + TERR + H2O
Key Takeaways

- Machine Learning creates analytic models to discover new insights, which humans cannot find.
- Event Processing uses analytic models - without redeveloping - to act on future events in real time.
- Machine Learning enables huge new opportunities for business.