From Prototype to Operative Software with RapidMiner

Data Analytics @ Lufthansa
Agenda

1. Lufthansa Industry Solutions: Who We Are

2. Our Daily Challenges

3. FlightPrediction @ Lufthansa

4. Lufthansa Industry Solutions & RapidMiner
We are Lufthansa Industry Solutions

IT Consulting and Systems Integration

Lufthansa Industry Solutions combines the dynamics of an SME with the economically secure background of Lufthansa; an internationally acting global corporation.
# Services overview

**Lufthansa Industry Solutions**

<table>
<thead>
<tr>
<th>Strategy and Management Consulting</th>
<th>IT Service Provider</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strategy Consulting</strong></td>
<td><strong>IT Service Provider</strong></td>
</tr>
<tr>
<td>Organization &amp; Process Consulting</td>
<td>IT System Operations</td>
</tr>
<tr>
<td>IT Consulting (processes, technologies, infrastructure)</td>
<td>Infra-structure</td>
</tr>
<tr>
<td>Process Organizational structure</td>
<td>Application Mgmt. &amp; Operations</td>
</tr>
<tr>
<td>Process Consulting / Process Organization</td>
<td>Operations of full business processes</td>
</tr>
</tbody>
</table>

**IT Consulting and System Integration**

**Lufthansa Industry Solutions**
Our Locations

- Norderstedt
- Berlin
- Wetzlar
- Hamburg
- Basel
- Cologne
- Oldenburg
- Miami
- Frankfurt
- Wolfsburg
Our Locations

Hard Facts

Managing Director
Bernd Appel

Founding
1998 as Industry Solutions division within Lufthansa Systems AG

100% owned by Lufthansa

208 Mil. € total revenue (46% of which within the Lufthansa Group)

>1300 trained employees

>200 customers
From Prototype to Operative - Data Analytics @ Lufthansa

Passenger Airlines
- Lufthansa Passenger
- Swiss
- Austrian

Aviation Industry
- Lufthansa Technik
- LSG Sky Chefs

Finance and Service Companies
- AirPlus
- LGBS

Air Freight
- Lufthansa Cargo

Digital Service Companies
- Lufthansa Industry Solutions
- Lufthansa Systems
Many Projects Successfully Delivered
About Me

Quick Facts

Fabian Werner
Born 1985
Ph.D. Mathematics
(Number theory)

Past

Present

Future

- Predictive analytics:
- Classification, regression, recommender systems, probabilistic modelling, ...
- RapidMiner, R, Python, PHP, Teradata, Oracle, Hadoop, ...

- Probabilistic modelling
- Time series analysis
- Bayesian learning
- Learning theory

\[ G_k(\tau) = 2\zeta(k) + 2 \frac{(2\pi i)^k}{(k - 1)!} \sum_{m=1}^{\infty} \sigma_{k-1}(m)e^{2\pi i m \tau} \]
## Our Daily Challenges

**Lufthansa Internal**

<table>
<thead>
<tr>
<th>Dashboarding</th>
<th>Predictive Analytics</th>
<th>(Deep) Machine Learning</th>
<th>Data Ingestion</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Dashboarding Icon" /></td>
<td><img src="image" alt="Predictive Analytics Icon" /></td>
<td><img src="image" alt="Machine Learning Icon" /></td>
<td><img src="image" alt="Data Ingestion Icon" /></td>
</tr>
</tbody>
</table>

- **Lufthansa**
  - Airplane punctuality
  - Competitor information
  - Airfreight load monitor

- **Lufthansa Cargo**
  - Arrival times
  - Passenger connection
  - Fuel & Weight forecast
  - Predictive maintenance
  - Recommend flights

- **Lufthansa Technik**
  - Document and image recognition
  - Damage recognition via audio data
  - Bots (see Mildred)
  - Live flight data
  - Weather data streams
  - Airplane sensor data
Process Model

Scoping Phase

Implementation stage

Laboratory stage

Use Case Driven Approach
Data Driven Approach

Business Analysis
Data Acquisition
Data Management
Data Science

Process and IT Design
Model Building

Process Improvement
Automation
Visualization
Operation
FlightPrediction @ Lufthansa
We predict the arrival time on aircraft takeoff
The Cycle
Goal & Process

Prediction Time

Prediction Point

Gate - to - Gate

Airtime

Rolling

Departure

Takeoff

Landing

Arrival
Goal & Process

Prediction Time

Prediction Point

Gate - to - Gate

Airtime

Rolling

Departure  Takeoff  Landing  Arrival
Goal & Process

From Prototype to Operative - Data Analytics @ Lufthansa

Accuracy of existing prediction
Goal & Process

Prediction Time

Prediction Point

Gate - to - Gate

Airtime

Rolling

Departure  Takeoff  Landing  Arrival

Accuracy of existing prediction
How? Unleashing The Power of Big Data

Historized Flight Data Stream

Eurocontrol

Weather Data Stream

Historic Data

Real-Time Data

\[ z_k = \beta_0 + \sum_{j=1}^{J} \beta_j \cdot x_{j,k} + u_k \]

Data Warehouse

rapidminer

Reporting Server

Predicted Arrival Time
The Cycle
Parameters for the Model

Flight Duration

- Delay @ Takeoff
- Season
- Lagged features
- Weather (en route & @destination)
- Route, distance, …
- Weight
- Traffic situation
- Runway(s)
- Passenger data
  - #Passengers with connection
  - % occupied seats
Parameters for the Model

- **Flight Duration**
  - Weight
  - Delay @ Takeoff
  - Season
  - Lagged features
  - Special disruptions
  - Weather (en route & @destination)
  - Actual data needed
  - Route, distance, …
  - Traffic situation
  - Capacity index & airport information still missing
  - Passenger data
    - #Passengers with connection
    - % occupied seats
  - Runway(s)
The Cycle
A few rolling window features are good!

Prefer models that can be trained fast: Play around with features instead of models.
The Cycle
Visualization: Performance Monitor

Data Made Up (No Actual Data)!
How To Save Money?

Gate Z50

Gate Z99
How To Save Money?
How To Save Money?
Visualization: Benefits Monitor

Data Made Up (No Actual Data)!
The Cycle
Automation: Supervisor
Supervision of Model

- Runs every hour
- Asserts that …
  - enough data
  - enough predictions
  - predictions produced fast
  - not too much missing data
  - …
Lessons Learned
Go To The Real World Fast – Make Everybody Understand The Benefits

- Supervisor & dashboard as early as possible
- Business needs to understand the real impact
- Translate Data Science to business language

- Need for standards
- Prototype needs to be translated to real application
- You do not want to rebuild everything
Partnership
Lufthansa Industry Solutions | RapidMiner

- Partnership founded in 2015
- RM widely used in Lufthansa
- Keeps up with market technologies
- For prototyping and productive applications
- Still lacks some software engineering capabilities …
- We pair it up with Hadoop, Oracle, Teradata, R, Python, …
- Lufthansa Industry Solutions: certified RapidMiner trainers
Many Thanks for Your Attention!