Machine Learning for Oil-Field Operations
Production Well Downtime Classification and Prediction

Madalina Trifan & Timothy Crossett
SAP Innovative Business Solutions
20 March 2019 | Stavanger
TOPICS

- If „One Size Fits All“ – Why Innovate?
- Energy sector project examples
- The IBSO Value Map of Products & Projects
- Data Science – Who We Are
- Machine Learning for Oil-Field Operations
- Production Well Downtime Classification and Prediction
SAP Innovative Business Solutions
Because One Size Doesn’t Necessarily Fit All

Goes beyond the SAP Standard portfolio.
Innovate quickly to respond to:
- Changing market conditions
- Your customer-specific requirements
- Integration of 3rd Party or In-House
- On-Premise / Cloud / Mobile

- Gives you First-Mover Advantage
- Built exactly to your needs
- Full support just like SAP Standard
Innovating For The Energy Sector.....

Broad Client Base....
Digitising Aircraft Refueling

- Into-Plane Refueling Solution
- Now live at over 130 airports
- Mobile capture of delivery docs
- Cancels paper-based processing
- Faster refueling turnaround
- Reduced anomalies
- Rapid billing through direct linkage to back-end

Featured in Forbes

Maintenance Technician Usability

Technicians were not recording cause-of-failure in a manner which could be used for trend analysis.

Solution created to simplify capturing details of failure reporting through PM notification.
Data-Science Analysis of Drone Data to Predict Tree-Growth Threats....

- Remote Sensing with focus on Vegetation Analysis
- Automated-classification
- Interpretation of different Geo Indexes for Vegetation monitoring
- Advanced predictive models to support decision algorithms
- Route construction based on ranking algorithms
- Route optimization for year over year work planning targeted to minimize risk
- Spot trimming recommendations for individual corridors.
- Built on SAP HANA
Addressing an Oil Major’s Retail Challenges

Mobile Loyalty….

- Mobile-First Loyalty Programme
- Rewards Customers Transactions
- Not Just Fuel-based Rewards
- Also Fill-Up-And-Go
01 | **Products**
Focused Business Solution (FBS)
Repeatable Custom Solution (RCS)
Listed in the regular SAP Pricelist with price metric (globally available)

02 | **Projects**
Applications previously built and delivered via a project

03 | **Topics**
Potentials SAP Innovative Business Solutions – not yet delivered
Deliverable as product or project
Oil and Gas Value map of SAP Innovative Business Solutions Products

- **Hydrocarbon Production**
  - Engineering
  - Hydrocarbon Production Management
  - Hydrocarbon Revenue Management
  - Field Logistics
  - Project Logistics

- **Hydrocarbon Logistics**
  - Commodity Management
  - Hydrocarbon Supply and Distribution
  - Commercial Process and Visibility
  - Secondary Distribution and Fuels Retailing

- **Operational Integrity**
  - Environment, Health, and Safety
  - Total Workforce Management
  - Asset Operations and Maintenance
  - Spare Parts and Material Management

- **Human Resources**
  - Core Human Resources and Service Delivery
  - Payroll and Time & Attendance Management
  - Talent Acquisition
  - People Analytics

- **Finance**
  - Accounts Payable and Receivable
  - Real Estate Management
  - Financial Planning and Analysis
  - Accounting and Financial Close
  - Finance Operations
  - Treasury Management
  - Enterprise, Risk & Compliance
  - Cybersecurity and Data Protection

- **Procurement**
  - Supplier Management
  - Inventory and Basic Warehouse Management
  - Service Procurement
  - Central Procurement
  - Invoice and Payables Management
  - Procurement Analytics
  - Supplier Collaboration

- **Digital Asset Management**
  - Intelligent Asset Management

- **Analytics**
  - Machine Learning Service
  - Business Intelligence
  - Planning and Analysis
  - Predictive Analytics

- **Application Platform and Infrastructure**
  - Application Platform
  - Interface Management
  - Process Management and Integration
  - Content and Collaboration

- **Data Management**
  - Database Management System
  - Data Warehousing
  - Big Data

- **IT Management**
  - IT Infrastructure Management
  - Application Lifecycle Management
  - Security Software
  - Identity and Access Management

- **IoT Business and Technology Service**
  - IoT Business Services
  - IoT Technology Service
  - IoT Edge

- **IOE, identity based performance processing**
  - GTS, identity based performance processing

- **Demand Signal Management**
  - Payroll and Time & Attendance Management

- **Success Factors Work Life**
  - Talent Acquisition

- **Asset Retirement Obligation**
  - People Analytics

- **IT Infrastructure Management**
  - Application and IT Infrastructure Security

- **Application Lifecycle Management**
  - Supplier Collaboration

- **Data Warehousing**

- **Big Data**
  - Application and IT Infrastructure Security

- **IoT Technology Service**
  - IoT Edge

- **IOE, identity based performance processing**
  - GTS, identity based performance processing

- **Multi-Resource Scheduling**

- **Multi-Bank Connectivity**

- **Treasury & Risk Mgmt, impairment acc. for exp. losses**

- **Transaction Avail. For Remote Sites**
  - Yard Logistics
  - Remote Logistics Management
  - Multi-Resource Scheduling

- **Transaction Avail. For Remote Sites**
  - Yard Logistics
  - Remote Logistics Management
  - Multi-Resource Scheduling

- **UI Logging**

- **Field Masking**

- **Micro Focus**

- **Adv. Data Synch. PW f. Mobile Asset Mgmt**

- **SAP Innovative Business Solutions**

- **Make Innovation Real**

© 2019 SAP SE or an SAP affiliate company. All rights reserved.
Who we are.

Focus Area Data Science & Machine Learning

Our group consists of 30 experts developing mathematical business software around the globe. With our skillset we bridge the gap between Business, IT and Science.

Standardized Approach to Enterprise Data Science

Our goal is to deliver enterprise software rather than scientifical reports. We leverage a standardized agile project approach and are relying on enterprise-proven toolsets.

References
Upstream Oil Company

Independent Exploration & Production company based in Texas

Focused on most significant oil-rich resource plays in the U.S.

Driving innovation with initiatives such as The Digital Oil Field
Artificial Lift
Most common lift types

Gas Lift

Rod Pump
Business Problem

- Monitoring >1500 wells with >100 measurement series each cannot be done manually.
- Often unplanned downtimes can only be treated re-actively.
- Preventive actions as well as pro-active planning would help to optimize resource usage and decrease revenue loss.
- Please note – this is an onshore use-case example, in offshore production operations such a solution which accurately predicts well downtimes will be equally imperative.
2 Major Downtime Reasons

Paraffin Build-up

Production pipe gets clogged by paraffin which builds up on the surface.

Well Loading

Mixture of Water, Oil and Gas gets so heavy that lifting stops.
Solution Approach

How to make oil fields smart.
Covered Aspects

1. Anomaly Detection
   Detect and flag anomalies in measurement data

2. Downtime Classification
   Classify the root cause for a current downtime. Reason codes include paraffin build-up and well loading issues.

3. Downtime Prediction
   Predict an upcoming downtime up to three days ahead. Root causes include paraffin build-up and well loading issues.
High-level approach

Field data arriving every 5 minutes in SAP HANA

Determining anomalies and applying models in real-time

Displaying result in Digital Oilfield Application

Technology

SAP HANA

Automated Predictive Library

SAP UI5
Anomaly Detection
**Anomaly Detection**

Mathematical Background (1/2)

Two kinds of anomalies:

1. Anomalies based on **absolute value**: For a specific measure a certain interval of values is expected. An anomaly may be a value outside this interval.

2. Anomalies based on **speed of change**: For a specific measure a certain speed of change is expected. An anomaly may be detected if the measurement changes faster or slower than expected.
Anomaly Detection
Mathematical Background (2/2)

- Current values are compared to historical average values
- The historical standard deviation will be used to determine the threshold
- Statistics will be updated asynchronously in batch mode
- For anomalies which are based on 'speed of change' the same logic will be applied to the derivative curve (delta between measurements)
Anomaly Detection
Configuration

Thresholds for pressure measurements may be different from thresholds for volume measurements. For this reasons the sensitivity parameters can be set for each measurement series differently.

<table>
<thead>
<tr>
<th>Measurement</th>
<th>Absolute Threshold</th>
<th>Derivative Threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td>Casing Pressure</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Water TVOL</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>

• Casing pressure may deviate up to 3 standard deviations from the average value
• We allow more 'peaks' in speed of change for casing pressur (5 std. dev.)
• We do not measure anomalies based on absolute value for Water TVOL (value 0 = ignore)
• Water TVOL will only be analyzed based on the derivative
• Measures which are not in the configuration list will be ignored
Downtime Classification
Downtime Classification

Overview

The Task

With Downtime Classification we try to find the root cause of a downtime at the moment it appears. This is different from prediction since we already know that there is a downtime at all.

The Benefit

For each downtime the models will be able to provide a score which indicates if well loading or paraffin build-up are probable root causes. This way production engineers can save time with analysis and quickly verify the suggested causes by looking at the indicative measurements.

Saving analysis time leads to faster reaction and thus to reduced downtime of the well.
Downtime Classification
Training the Model

Training process

Detect Downtime
- Based on gas flow rate
- Gas flow below 5 for at least 2 hours

Find reason in comments
- Join comments with timely correlation
- Search for keywords such as ‘paraffin’ or ‘well loaded up’

Engineer features
- Gather historical and current measurements
- Engineer features such as water cut from the data

Train model
- Train classification model based on historical downtimes
- One model per root cause

Application process

Detect Downtime

Engineer features

Apply model
- Apply models that have been trained before
- Generate score for each trained root cause
Downtime Classification

Well loading

Predictive Power: \(0.8492\)
Prediction Confidence: \(0.9822\)

We need to look at 6.5% of the downtimes in order to capture at least 70% of the well loading downtimes.
Downtime Classification

Paraffin build-up

Predictive Power: 0.8148
Prediction Confidence: 0.9725

We need to look at **7.7%** of the downtimes in order to capture at least **70%** of the paraffin build-up downtimes.
Downtime Classification
Summary

• An **hourly report** gets triggered showing all **current downtimes**

• By using the classification scores 70% of the well loading and paraffin downtimes can be detected by **looking at less than 8% of the downtimes**

• The report will **reduce reaction times** and decrease downtimes
Downtime Prediction
Downtime prediction considers three concepts:

1. **Historical data** separated by short, mid and long-term influencers
2. **Blackout period** to prevent that the prediction ultimately turns into classification
3. **Prediction period**. In this period the algorithm tries to predict downtimes.

The above picture refers to one record which is used for training the prediction. One of those records will be generated for each well and each quarter of the day.

1500 wells x 365 days x 4 quarters per day ~ **2.2 million** training records per year
Downtime Prediction

Data Preparation

Considered **measures per record:**

<table>
<thead>
<tr>
<th>Measure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>WELL_TYPE</td>
<td>Free flow, Gas lift or Rod pump</td>
</tr>
<tr>
<td>PROD.HOUR</td>
<td>Production Time</td>
</tr>
<tr>
<td>DW_SIGN</td>
<td></td>
</tr>
<tr>
<td>CASING_PRESS</td>
<td>Casing Pressure</td>
</tr>
<tr>
<td>TUBING_PRESS</td>
<td>Tubing Pressure</td>
</tr>
<tr>
<td>SEP_PRESS</td>
<td>Seperator Pressure</td>
</tr>
<tr>
<td>GASLIFT_STATIC_PRESSURE</td>
<td>Gaslift static pressure</td>
</tr>
<tr>
<td>CORI_METER_DENSITY</td>
<td>Coriolis meter density</td>
</tr>
<tr>
<td>CORI_METER_TEMP</td>
<td>Coriolis meter temperature</td>
</tr>
<tr>
<td>PUMP_FILLAGE</td>
<td></td>
</tr>
<tr>
<td>STROKES_PER_MINUTE</td>
<td>Stroke per minute for Rod Pump</td>
</tr>
<tr>
<td>LOAD_SPAN</td>
<td></td>
</tr>
<tr>
<td>PUMP_HP</td>
<td></td>
</tr>
<tr>
<td>WELL_DOWN_STATUS_CALC</td>
<td>Calculated status field for downtime</td>
</tr>
<tr>
<td>GASLIFT_TEMP</td>
<td>Gaslift temperature</td>
</tr>
<tr>
<td>PID_VALVE_POS_FB</td>
<td></td>
</tr>
<tr>
<td>PID_MANUAL_SP</td>
<td></td>
</tr>
<tr>
<td>PID_SELECTOR</td>
<td></td>
</tr>
<tr>
<td>PID_AUTO_SP</td>
<td></td>
</tr>
<tr>
<td>PID_VALVE_OUTPUT</td>
<td></td>
</tr>
</tbody>
</table>

Simple **derived fields:**

<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>WATER_CUT</td>
<td>Ratio of water vs. Oil</td>
</tr>
<tr>
<td>PRESS_DIFF_1</td>
<td>Diff. between casing and tubing pressure</td>
</tr>
<tr>
<td>PRESS_DIFF_2</td>
<td>Diff. between casing and gaslift pressure</td>
</tr>
<tr>
<td>SEC_SINCE_LAST_WL</td>
<td>Time since last well loading downtime</td>
</tr>
<tr>
<td>SEC_SINCE_LAST_PF</td>
<td>Time since last paraffin build up downtime</td>
</tr>
<tr>
<td>SEC_SINCE_LAST_DT</td>
<td>Time since last downtime</td>
</tr>
<tr>
<td>LIFE_WELL</td>
<td>Duration of overall well life</td>
</tr>
<tr>
<td>LIFE_WELLTYPE</td>
<td>Duration of well life in current production mode</td>
</tr>
<tr>
<td>DT_WL</td>
<td>Flag if there currently is a well loading downtime</td>
</tr>
<tr>
<td>DT_PF</td>
<td>Flag if there currently is a paraffin related downtime</td>
</tr>
<tr>
<td>DT</td>
<td>Flag if there currently is a downtime</td>
</tr>
</tbody>
</table>

Complex **derived fields:**

<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggregates</td>
<td>For each of the other fields AVG, MIN and MAX values are calculated for each period (short, mid and long-term)</td>
</tr>
<tr>
<td>Derivatives</td>
<td>For each of the other fields the ratio between the short term and the long-term period is calculated</td>
</tr>
<tr>
<td>Deviations</td>
<td>For some of the fields the deviation of MIN and MAX values against AVG are calculated</td>
</tr>
</tbody>
</table>

Overall more than **350 features** are taken into account!
Downtime Prediction
Dataset used for Training

Consider 2 years of training data

Training of prediction model and automatic feature selection lasts several hours

Feature generation and model inference runs in real-time!

350 features
4 million records
1.4 billion cells to evaluate
Downtime Prediction

Well loading

Predictive Power: 0.7923
Prediction Confidence: 0.9622

Around **30 features** have been identified of being relevant.
Downtime Prediction

Paraffin build-up

Predictive Power: 0.8106
Prediction Confidence: 0.9746

Around **40 features** have been identified of being relevant.
Downtime Prediction
Summary

Key **takeaways** from the statistical analysis:

- The prediction models are **stable**, which means we expect the same results on production data.
- The model have a high predictive power, which means that there is a certain **predictability** for the given events.
- A majority of the variable contributions, which were determined with a quantitative approach, could be verified by **known physical relations**.

*Water Cut* has been found relevant for well loading whereas *Coriolis temperature* had a high influence on Paraffin Build-up. This fact creates **additional confidence** in the models.
Mobile **User Interface**

Getting the insight on-site.
Demo

Mobile Frontend for Field Engineers
Wrap Up
Key Takeaways

- SAP Innovative Business Solutions enhanced the Digital Oil Field Platform of our client by a Machine Learning component.

- This enables the client:
  - to detect anomalies on measurement data,
  - classify historical downtimes and
  - predict upcoming downtimes up to 3 days into the future.

- The solution includes a mobile user interface for Field Engineers.
Next steps

• Expand the solution for **2 more generalized** types of downtimes:
  • solids accumulation
  • flumping well
• Apply the solution to **2 more production fields** (with more than 1000 wells each).
• Tank level analysis and forecasting for oil fields with **no well production meters**.
• Forecast well production to better schedule tanks draining.
• Cleansed and projected oil flow to **serve as input** for the prediction of downtimes.
Thank you!

Contact information:

**Madalina Trifan**  
Senior Data Scientist  
SAP Innovative Business Solutions  
[madalina.trifan@sap.com](mailto:madalina.trifan@sap.com)  
+40 756040485

**Timothy Crossett**  
Business Development Manager  
SAP Innovative Business Solutions  
[timothy.crossett@sap.com](mailto:timothy.crossett@sap.com)  
+49 1516 2345 442