Follow-Up Project: Automated Text Mining on Job Offers Using SAP HANA
Analyzing Skill and Competency Requirements for Industry 4.0

Situation

Industrie 4.0/Industry 4.0/the Industrial Internet: combining known technologies in a new way: disruptive solutions
- The Internet of Things
- Cyber Physical Systems (CPS)
- Smart Factories
- Embedded Systems

(Kagermann et al. 2013, Industrie 4.0 2016)

Complication

The way we live and work will change significantly
- New ways of business value creation
- Changing business models and strategies
- Adjusted business processes

(Kagermann et al. 2013)

Resolution

Apply text analysis and text mining offered by SAP HANA and the SAP Predictive Analysis Library (PAL) on job offers collected from German online job portals to extract skill and competency requirements for Industry 4.0:

- Build on technical experience from former projects
- Apply technical knowledge to a new area
- Try to discover competency profiles from analyzing job offers

System

- SAP HANA 2 SPS01, 1 TB RAM, 32 Cores (CPU)
- SAP HANA Studio, version 2.3.10
- PAL library
- Python, version 3.5.2
- Node.js 6.9.5 LTS

Data Sets

- Manually collected German job offers, most tests done on collection of Nov – Apr. (2,470 job offers)
- Test data sets: T1: 15 job offers, T2: 50 job offers, manually classified

<table>
<thead>
<tr>
<th>Month</th>
<th>Portal A</th>
<th>Portal B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nov 2016</td>
<td>176</td>
<td>160</td>
</tr>
<tr>
<td>Dec 2016</td>
<td>271</td>
<td>250</td>
</tr>
<tr>
<td>Jan 2017</td>
<td>310</td>
<td>375</td>
</tr>
<tr>
<td>Feb 2017</td>
<td>251</td>
<td>174</td>
</tr>
<tr>
<td>Mar 2017</td>
<td>175</td>
<td>176</td>
</tr>
<tr>
<td>Apr 2017</td>
<td>273</td>
<td>198</td>
</tr>
<tr>
<td>May 2017</td>
<td>251</td>
<td>176</td>
</tr>
<tr>
<td>Jun 2017</td>
<td>251</td>
<td>175</td>
</tr>
<tr>
<td>Total</td>
<td>1,307</td>
<td>1,225</td>
</tr>
</tbody>
</table>

Next Steps

- Integrate web crawler into SAP HANA application
- Extend functionality of web crawler
- Improve job profile discovery
- Analysis of skill and competency requirements with regards to metadata (e.g., job title, company, region)
- Implementation of a regular analysis to discover changes over time