What’s new...

• 2+1 feature releases last year: 2.12, (3.0), 3.1 (only KNIME Analytics Platform + Server)

• Changes documented online...
What’s new pages and YouTube

What is new in KNIME Analytics Platform 3.1 and in KNIME Server 4.2

Published on Dec 6, 2015

This video illustrates the main new features in KNIME Analytics Platform 3.1 and KNIME Server 4.2.

The KNIME Analytics Platform 3.1 has a new Look & Feel, including a new look for node icons and for...
Changelog...

KNIME Analytic Platform v2.13.0

© 2016 KNIME.com AG. All Rights Reserved.
Changelog ...

New nodes/features by version:

<table>
<thead>
<tr>
<th>Version</th>
<th># new nodes / sets</th>
<th># features</th>
</tr>
</thead>
<tbody>
<tr>
<td>v2.12</td>
<td>27</td>
<td>53</td>
</tr>
<tr>
<td>v3.0</td>
<td>3</td>
<td>28</td>
</tr>
<tr>
<td>v3.1</td>
<td>14</td>
<td>21</td>
</tr>
</tbody>
</table>
Outline

Interactive feature demos
  ... by the developers
Outline

• Analytics / Mining
• Statistics
• PMML - Standardizing predictive models
• Streaming Executor

• KNIME Server
• KNIME Productivity Extension: Workflow Diff
Analytics / Mining
Analytics / Mining

• Trees / Forest / Ensembles

• Active Learning

• R Integration
Trees / Forest / Ensembles

• Demo
Active Learning

- Labs Extension
- Involve user to construct training data set
- Workflow loop to query and label ‘interesting’ data points
- Used user-labeled data set on remaining data
Active Learning (example from Node.Pedia)
R Integration

• Rewrite of infrastructure
  – Significantly faster
  – Concurrent execution

• No change of usage model
Statistics

Christian Dietz
Statistics Nodes

What is it about?

• Several new useful statistic nodes in KNIME.
• Thanks to Bob Muenchen (University of Tennessee).
• Work in progress! We are still adding nodes.
Statistics Nodes

Parametric and nonparametric hypothesis testing

Linear Discriminant Analysis
Statistics Nodes

Cronbach Alpha

Rank Correlation
Statistics Nodes

Odds and Risk Ratios

And more to come...
Statistics Nodes

ARIMA for time series analysis

- “Auto regressive integrated moving average”
- Several ARIMA related nodes in KNIME
PMML - Standardizing Predictive Models

Alexander Fillbrunn
What is PMML?

• Predictive Model Markup Language
• XML based standard for predictive models
• KNIME can export most of its models as PMML
• To consume 3rd-party models, a scoring engine such as Zementis Adapa/UPPI is more suitable
PMML Creation in KNIME

• Special port for PMML models
• Supported by most KNIME learners
  – Decision Trees, Neural Nets, ...
  – Ensembles
• Also used for Preprocessing
  – Normalizing, Binning, Missing Values, ...
• Modular PMML
  – Built step by step parallel to the data flow
Demo: Modular PMML
Decision Tree to Ruleset

- Transforms a decision tree to a PMML ruleset model
  - Easier to interpret
- Also outputs rules as a KNIME table
  - Easier to export & deploy
  - Can be manipulated using standard KNIME nodes
Applying Rule sets

• New node: Rule Engine (Dictionary)
  – Input: data and ruleset table
  – Output: Results and optional PMML model

• Import rules from other sources
• Mix rules from multiple sources
Streaming Executor

Martin Horn
Streaming

• Default Execution
Streaming

• Streaming Execution
Streaming

• Row-wise
• Process, pass & forget → Faster with less I/O overhead
• Concurrent execution
Streaming

• Demo
Streaming - Demo

Diagram showing the preprocessing steps for sentiment classification in KNIME.
Streaming - Demo
Streaming - Demo
Streaming - Demo

*0: StreamedSentimentClassification

[Diagram of a KNIME workflow showing file reader, partitioning, and a dialog box for settings such as job manager selection and sentiment color management.]
Streaming - Demo
Streaming – Pros and Cons

Advantages

• Less I/O overhead (process, pass & forget)
• Parallelization

Disadvantages

• No intermediate results, no interactive execution
• Not all nodes can be streamed
Streaming – Streamed Nodes

- More than 100 Nodes
- Text Processing Nodes
- Image Processing Nodes
- ...
Streaming – How-to?

1. Streaming Execution Plugin (Labs) limitations: loops, metanodes

2. Wrapped Metanode

3. Set ‘Simple Streaming’-Job Manager
KNIME Server ... and related

Thorsten Meinl
Overview

• MongoDB and JSON
• WebPortal templates
• Advanced scheduled execution
• RESTful webservice interface
• Automation by calling other workflows
MongoDB and JSON (I)

- MongoDB is a NoSQL database based on JSON
- Special set of nodes
  - due to lack of a standard SQL interface
MongoDB and JSON (II)

- JSON nodes for working with JSON data
  - Similar to the XML nodes
- Use combination of MongoDB and JSON nodes

![Diagram showing flow of data processing with MongoDB Reader, JSON Path, Color Manager, Scatter Plot, Table Creator, String to JSON, and MongoDB Update nodes.](image-url)
WebPortal templates (I)

• Layout can be configured by templates
  – Footer & header
  – Main panel
  – Login page
  – ...

• Custom stylesheet and common JS libraries

• Custom templates are part of the configuration and are not overridden by server updates
WebPortal templates (II)
Advanced scheduled execution (I)

• Requires KNIME Server(Space) 4.2
• Allows setting filters for repeating jobs
  – By day-of-week
  – By day-of-month
  – By month
• Skip execution if previous job is still running (optional)
• Respects daylight saving for daily schedules
• Scheduled jobs can be edited and disabled
Advanced scheduled execution (II)
RESTful webservice interface (I)

• Main addition to KNIME Server 4.1
• REST = Representational State Transfer
  – Communication based on HTTP
  – Usually clear text (JSON, XML, ...)
• Many possible clients
  – Web browser
  – Java applications (e.g. via JAX-RS)
  – KREST nodes :-)
• Goal: complete server interface based on REST
RESTful webservice interface (II)

• Functionality currently exposed via REST
  – Repository browsing
  – Executing jobs
    • With optional input parameters
    • Synchronously, asynchronously
    • Retrieving results and/or report
  – Show permissions
  – Show all running jobs
• Read our corresponding blog posts for more details
Automation by calling other workflows (I)

- Call Remote/Local Workflow nodes
- Called workflow parameterized via quickform nodes
  - In principle every quickform node can be used for a REST-enabled workflow
Automation by calling other workflows (II)

• This is how the JSON looks like...

```
{
    ...
    "hasReport" : false,
    "inputParameters" : {
        "model-number-20" : {
            "integer" : 0
        },
        "data-record-1" : {
            "Col0" : "A11",
            "Col1" : 6,
            ...
            "Col18" : "A192",
            "Col19" : "A201"
        }
    },
    "workflow" : "/Summit/workflows/Predictor Flow",
    ...
}
```
Automation by calling other workflows (III)

- Calling a remote workflow
Automation by calling other workflows (IV)

- Also works for local workflows via the Call Local Workflow node
  - Part of the Personal Productivity Extensions
Productivity Extensions: Workflow Diff

Ferry Abt
KNIME WorkflowDiff

Snapshot in /Events/2016_02_Summit/Whats new/WorkflowDiff (Ferry)

<table>
<thead>
<tr>
<th>Element/Creation date</th>
<th>Creator</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>WorkflowDiff</td>
<td></td>
<td>sampling &lt;-&gt; Filter</td>
</tr>
<tr>
<td>Thu 2016-02-18, 17:07:03h</td>
<td>ferry.abt</td>
<td>Filter for P.Cood.</td>
</tr>
<tr>
<td>Thu 2016-02-18, 17:06:14h</td>
<td>ferry.abt</td>
<td>Color</td>
</tr>
<tr>
<td>Thu 2016-02-18, 17:05:26h</td>
<td>ferry.abt</td>
<td></td>
</tr>
<tr>
<td>Thu 2016-02-18, 17:04:58h</td>
<td>ferry.abt</td>
<td>original</td>
</tr>
</tbody>
</table>

Download snapshot

Compare
KNIME WorkflowDiff
KNIME WorkflowDiff
KNIME WorkflowDiff
KNIME WorkflowDiff
Workflow Diff

• Demo
Workflow Diff - Recap

• Identifies changes in workflow-structure
• Aligns workflows to identify differences
• Available for KNIME Server and KNIME Productivity Extensions
Summary

• Questions / Interested in demo / comments?
  → Talk to us in the breaks / at the booth

• Interested in What’s cooking?
  → After lunch

• Interesting in Big Data?
  → After lunch
The KNIME® trademark and logo and OPEN FOR INNOVATION® trademark are used by KNIME.com AG under license from KNIME GmbH, and are registered in the United States.