

Cheat Sheet: Aufbau eines KNIME Workflows für Anfänger

Erste Schritte mit der KNIME Analytics Platform

- Verwenden Sie den Getting Started Guide, um Ihre ersten Schritte mit visuellen Arbeitsabläufen bei www.knime.com/getting-started-guide.
- Öffnen Sie mehr über enthaltene Knoten und erkunden Sie Arbeitsbeispiele in der **KNIME Analytics Platform Version 5 Starter Perspective Collection** auf [KNIME Community Hub](#).

Node Action Bar: Direkt mit dem Knoten interagieren z.B. Konfigurieren, Ausführen, Löschen oder Zurücksetzen eines Knotens. Konfigurieren: Öffnen Sie den Konfigurationsdialog. Ausführung: Führt den Knoten aus. Abbrechen: Löscht die Ausführung des Knotens. Zurücksetzen: Setzt den Knoten zurück.

Node Labels: Doppelklick „Ein Kommentar hinzufügen“ unten unter dem Knoten, um einen Kommentar/Label hinzuzufügen. Dynamische Anschlüsse: Zusätzliche Eingänge können hinzugefügt, indem Sie auf das Plus links klicken Seite des Knotens.

Nicht konfiguriert: Node ist noch nicht konfiguriert und kann nicht mit seinen aktuellen Einstellungen ausgeführt.

Konfiguriert: Node wurde korrekt konfiguriert und kann jederzeit ausgeführt.

Ausgeführt: Node wurde erfolgreich durchgeführt und Ergebnisse kann in nachgeschalteten Knoten betrachtet und verwendet werden.

Fehler: Der Knoten hat bei der Ausführung einen Fehler aufgetreten.

ANALYZE

Scatter Plot: Vertreten Datenzeilen als Punkte eingegeben in einem zweidimensionalen Grundstück. Eingangsabmessungen (Spalten) auf der x-y-Achse Plot und Grafik Eigenschaften können geändert werden in der Konfiguration Fenster oder interaktiv in die Knotenansicht.

Sunburst Chart: Displays kategorische Säulen durch eine Hierarchie Ringe. Jeder Ring wird geschnitten nach dem Nominal Werte im Korrespondenten in Spalte und ausgewählter Hierarchie. Das ist ein mächtiges Diagramm für multivariate Analyse.

Gestapelte Fläche: Abbildung mehrere numerische Daten Spalten auf jeder andere Verwendung der vorherigen als Basisreferenz. Die Gebiete zwischen Linien sind für einfacher Vergleich. Dieses Diagramm ist häufig verwendet visualisieren trendige Themen.

Line Plot: Zeigt Zahlenwerte in Datenspalten (y-Achse) gegen Werte in einer Referenzspalte (x-Achse). Datenpunkte werden über farbige Leitungen verbunden, wenn die Referenzspalte auf der x-Achse enthält sortierte Zeitwerte, das Liniendiagramm graphisch stellt die Entwicklung einer Zeitreihe dar.

Farbmanagement: Jede Eingabezeile gibt eine Farbeigenschaft zu basierend auf dem Wert der Zeile in einer ausgewählten Spalte. Diese Farbeigenschaft beeinflusst die Grafik Darstellung in den kommenden Ansichten.

Daten Explorer: Bietet einen interaktiven Blick die Statistik der Eingabedaten zusammenfassen über statistische Maßnahmen und Histogramme - für numerische und nominale Spalten.

Box Plot: Visualisiert numerische Spalten mit dem Quartil Statistiken. Achten Sie auf die Punkte am Ende von den Whiskers - sie könnten Ausreißer markieren!

Interactive Pie Chart: Visualisiert eine aggregierte Metrik für verschiedene Daten Trennwände mit farbigen Scheiben auf einem Kreis, wo die Flächen sind proportional zu den metrischen Werten. Die Partitionen werden durch eine kategoriale Spalte definiert.

Balkendiagramm: Visualisiert eine oder mehrere aggregierte Metriken für verschiedene Datentrennwände mit rechteckigen Balken, wo die Höhen sind proportional zu den metrischen Werten. Die Partitionen werden durch eine kategoriale Spalte definiert.

DEPOLEN

LERNENDE NODEN

Lies CSV-Dateien. Es hat eine Auto-Detect-Funktion automatisch die Datei erraten Struktur. Wie bei anderen Lesern Knoten, Sie können ein "Dateisystem hinzufügen Anschluss" Eingangsport zum Anschluss zu verschiedenen Datenquellen.

Lesen Sie maschinelles Lernen Modelle mit jeder der Lernknoten. Modelle werden in der Regel gespeichert nach Training und wieder verwendet in der Bereitstellung.

Lesen Sie Daten von einem .table Datei..table Dateien sind organisiert mit einem KNIME proprietäres Format, einschließlich der vollständigen Datei Struktur und optimiert für Raum und Geschwindigkeit - Bereitstellung Maximale Leistung mit mindestens Konfiguration!

liest Daten von einem Google Sheet Datei. Authentication erfolgt auf der Website von Google. Google Anmeldeinformationen sind nicht innerhalb der KNIME Workflow.

In Lese- und Schreibknoten wird der Dateipfad relativ zu einem Schlüsselort der aktuelle KNIME-Installation, wie Workflow, Workflow-Datenbereich und Mountpoint.

PREDIKTOR NODEN

Entdecken Sie Lernen Predictor

TRANSFOS

Weiterlesen

ANALYSE

Bereitstellung

ANALYZE

Markiert die zu exportierende Datentabelle zu BIRT - a teilweise Open Source Reporting Tool integriert innerhalb von KNIME. Beim Wechsel von KNIME zu BIRT werden die markierten Datensätze in BIRT importiert. Der Knoten Image To Report markiert die Eingabebilder nach BIRT exportiert werden.

Predictor Nodes: Wird verwendet Anwendung von Modellen. Die beiden Eingänge sind das geschulte Modell und Daten zu verarbeiten. Der Ausgang die ursprünglichen Daten enthält und die Modellvorhersagen.

Daten zum Bericht

Excel Schreiber

Tabelle Schreiber

CSV Schreiber

Google Plauschreiber

an Tableau Server senden

ROC Kurve

Integrationen zu vielen Open Source-Datenanalyse-Tools sind auch verfügbar. Einige verwenden den KNIME-Knoten GUI (H2O, Weka, Keras, Spark MLlib). Andere bieten Knoten mit einem Entwicklungsumgebung für Skript und Debugging (R, Python, Java).

VERKEHR

Gruppe

Gruppen die Zeilen einer Tabelle durch die eindeutigen Werte in ausgewählten Spalten und berechnet Aggregation und statistische Maßnahmen für die definierten Gruppen. Trotz seines einfachen Namens, es bietet mächtig Funktionalität und hat viele nicht geprüft Verwendungen. Zum Beispiel - Zeileneduplikation.

Pivot

Erweitert die Aggregationsfunktionalität der Gruppe By node by create an Ausgabedaten Tabelle mit Spalten und Zeilen für das einzigartige Werte in ausgewählten Eingangsspalten. Anmerkung: eindeutige Werte der Gruppierungsspalte werden Zeilen und die eindeutigen Werte der Die Schwenksäule wird zu Säulen.

Regel Motor

Gilt für jede Zeile der Eingabe eine Reihe von Regeln Datentabelle. Alle Regel Engine Operatoren sind auch erhältlich im Column Expressions-Knoten.

Partitionierung

Daten in zwei Teilmengen nach einem Stichprobenstrategie. Dieser Knoten ist in der Regel zur Herstellung eines Trainings und eines Testsets ein Maschinenlernmodell ausbilden und bewerten.

Row Filter

Filtert Zeilen in oder aus den Eingabedaten Tabelle nach einer Filterregel. Die Filterregel kann einen Wert in einer ausgewählten Spalte oder Zahlen in einem numerischen Bereich.

Mathematik

Implementiert eine Reihe von Mathe-Operationen über mehrere Eingangsspalten, von einfach Summe und Durchschnitt, zu Logarithmen und exponentiellen. Alle Mathe Formel Operatoren sind auch im Column Expressions Knoten verfügbar.

String to Date&Time

Konvertiert Werte in einer Streichspalte in Datum und Uhrzeit. Datum und Uhrzeit in den String-Werten enthalten kann manuell definiert oder auto erkannt.

Zeilenteilung

Spaltet Werte in einer ausgewählten Spalte in zwei oder mehr Teilstrings, wie durch einen Abgrenzer definiert Match. Delimiter ist ein gesetztes Zeichen, wie z.B. ein Komma, Raum oder andere Charaktere oder Zeichenfolge.

Spaltenfilter

Filtert Spalten in oder aus den Eingabedaten Tabelle nach einer Filterregel. Spalten zu kann manuell abgeholt oder ausgewählt werden nach ihrer Art oder einem Regex Ausdruck, der ihrem Namen entspricht.

Umbenennung

Gibt neuen Namen und Typen zu ausgewählten Spalten, wie im Dialog konfiguriert.

Mitglied

Registriert Zeilen aus zwei Datentabellen basierend auf gemeinsame Werte in einem oder mehreren Schlüsseln Spalten. Der Ausgang - innen, links außen verbinden, rechts außen verbinden, voll außen verbinden, oder jeweilige Antijoins - kann in mehrere Ausgabetabellen.

Sortierung

Sortiert den Tisch im Auf- oder Absteigen nach den Werten eines gewählten Spalte. Außerdem ist es möglich, basierend auf mehreren Spalten.

Concaten

Verschmelzung von zwei oder mehr Datentabellen vertikal durch Aufpeilen von Zeilen in Spalten mit denselben Namen. Zeilen in nicht überlappenden Spalten werden mit fehlenden Werten gefüllt.

Fehlender Wert

Definiert eine Strategie, mit fehlenden umgehen Werte in der Eingabedaten - entweder global auf allen Spalten, oder einzeln für jede einzelne Spalte.

String Manipulation

Durchführung von Operationen auf String-Werten in Spalten, wie z.B. Kombination von zwei oder mehr Strings zusammen, ein oder mehrere Substrings, Beschrneiden von Leerräumen, und so auf. Alle Betreiber sind auch in der Column Expressions Node.

E-Learning Kurse

Nehmen Sie unsere kostenlose online selbstbefriedigte Kurse, um über die verschiedene Schritte in einem Data Science-Projekt (mit Übungen und Lösungen, um Ihre Wissen) bei [knime.com/knime-self-paced-Gänge](#)

KNIME Community Hub

Durchsuchen und Workflows, Knoten und Komponenten teilen. Bewertungen oder Kommentare zu anderen hinzufügen Arbeitsabläufe bei [wohnzimmer.de](#)

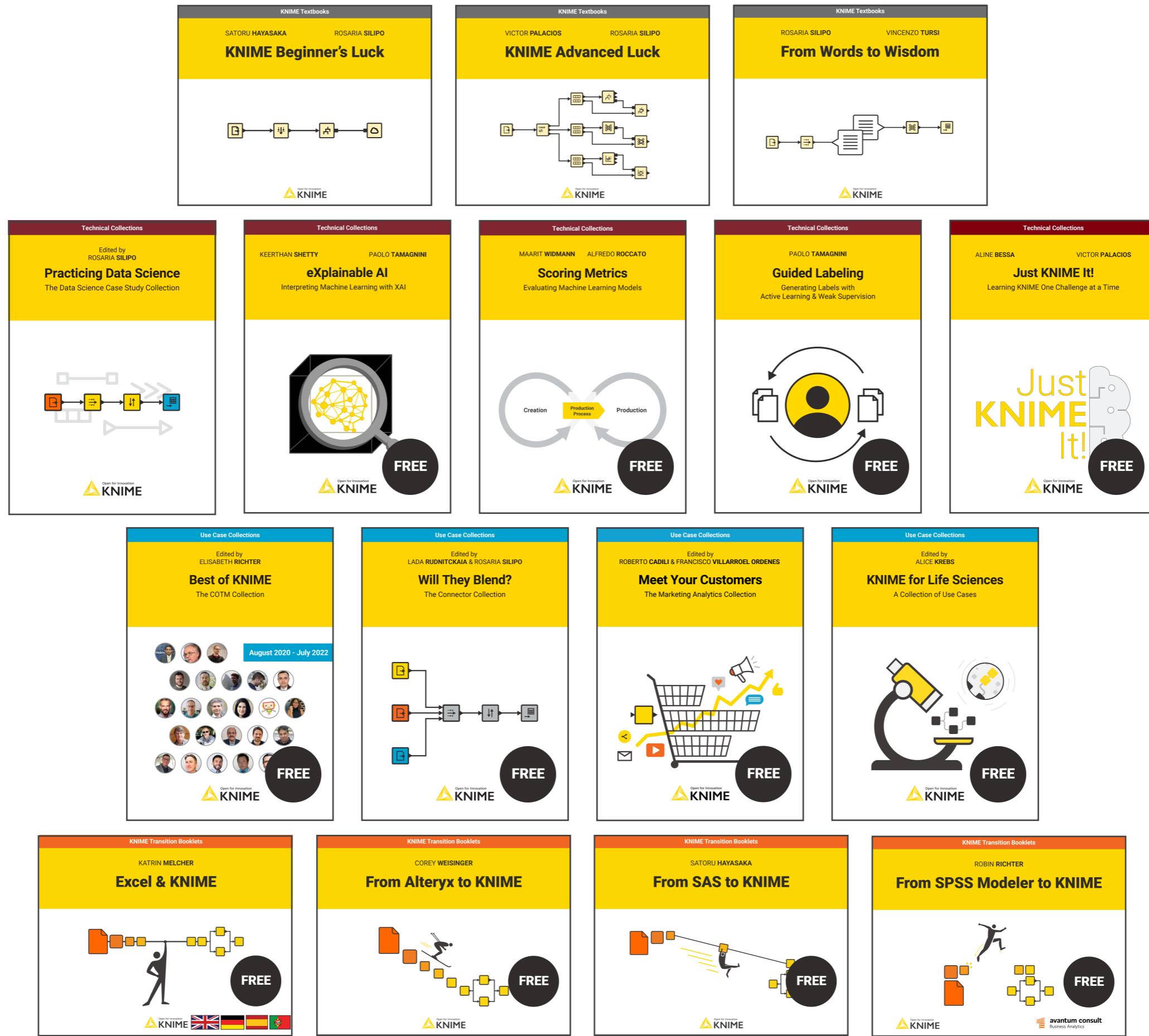
KNIME Forum

Mitglied unserer globalen Gemeinschaft und in Gesprächen bei [Forum.de](#)

KNIME Business Hub

Für Teams Zusammenarbeit, Automatisierung, Management, & Einsatz Check-in KNIME Unternehmen Hub bei [knime.com/knime-business-hub](#)

Erweitern Sie Ihr KNIME Wissen mit unserer Sammlung von Büchern von KNIME Press. Für Anfänger und Fortgeschrittene, bis hin zu denen, die an Fachthemen wie Themenerkennung, Datenvermischung und Klassik interessiert sind Lösungen für gängige Anwendungsfälle mit der KNIME Analytics Platform - es gibt für jeden etwas. Verfügbar unter www.knime.com/knimepress.



Brauchen Sie Hilfe?
Kontaktieren Sie uns!

