**Handout: KNIME Workflow-Dokumentation**

**1. Einführung: KNIME-Knoten und Datenverarbeitung**

* **CSV Reader: Daten Einlesen (Datenset\_G1.csv):**
  + Einlesen des Datensets aus der Datei "Datenset\_G1.csv".
* **Data Explorer: Erste Daten Exploration:**
  + Erste Dateneinsicht für statistische Informationen und Datenüberblick.
* **Column Remaner, Rule Engine:**
  + Übersetzung von Spalten und Hinzufügen neuer Spalten.
* **Row und Column Filter:**
  + Trennung der Daten in Über- und Unter-50K-Gruppen.
* **CSV Writer: CSV für PowerBI:**
  + Export der bearbeiteten Daten für PowerBI-Analyse.

**2. Zweite Dateneinsicht und Explorative Datenanalyse (EDA)**

* **Data Explorer: Zweite Daten Exploration:**
  + Weiterführende Analyse für tieferes Verständnis.
* **Linear Correlation, Value Counter, Heatmap, Bar Chart:**
  + Untersuchung von Zusammenhängen, Werteverteilung und Visualisierungen.

**3. Maschinelles Lernen: Modelle und Parameteroptimierung**

* **K Nearest Neighbor, Random Forest, Gradient Boosted Trees, Naive Bayes:**
  + Anwendung verschiedener Modelle auf die Daten.
* **Parameteroptimierung, Partitionierung, Scorer:**
  + Feinabstimmung der Modelle für optimale Leistung.

**4. Ergebnisse Transformieren: Modellbewertung und Vergleich**

* **Row und Column Filter, Constant Value Column:**
  + Filtern und Identifizieren von relevanten Informationen.
* **GroupBy: Mittelwert und Abweichung:**
  + Aggregation der Ergebnisse für eine bessere Analyse.

**5. Baseline-Analyse: Vergleichsmaßstab für Modelle**

* **Value Counter, Math Formula: Baseline Mehrheit/Gesamt:**
  + Berechnung der Baseline-Genauigkeit.
* **Vergleich mit Modellen:**
  + Integration der Baseline-Ergebnisse für Vergleichszwecke.

**6. Ergebnisse Zusammenführen und Speichern**

* **Concatenate, Joiner, Table View:**
  + Zusammenführung der Modellergebnisse für Gesamtübersicht.
* **CSV Writer: Modellergebnisse in CSV-Datei speichern:**
  + Speichern der aufbereiteten Ergebnisse für externe Nutzung.

**7. Ergebnisse überprüfen und analysieren**

* **CSV Reader: Modellergebnisse anzeigen:**
  + Einlesen und Überprüfen der gespeicherten Ergebnisse.
* **Data Explorer: Modellergebnisse Übersicht:**
  + Übersichtliche Darstellung der rohen Modellergebnisse.

**8. Ergebnisse aufbereiten und finalisieren**

* **Round Double, Column Filter, Column Manager:**
  + Aufbereitung und Strukturierung der Daten.
* **Timestamp Generator: Zeitstempel:**
  + Hinzufügen eines Zeitstempels für zeitliche Referenz.

**9. Fazit und Ausblick**

Die dokumentierte KNIME-Workflow umfasst Datenimport, -transformation und -analyse, die Anwendung verschiedener ML-Modelle und die umfassende Aufbereitung der Ergebnisse. Der Workflow bietet eine strukturierte und automatisierte Methode zur Datenauswertung und Modellbewertung.