Data Thinking & Canvas

Aufgabenblatt

Data Thinking & Prozessmodelle (CRISP-DM)

Lern-Story: Bevor Daten analysiert werden, muss man das *richtige Problem* stellen: Viele Data-Science-Projekte scheitern (60–80 %) an unklaren Zielen und fehlender Strategie, nicht an der Technik. *Data Thinking* stellt das Geschäftsproblem, die Beteiligten und den Wert in den Mittelpunkt — ein *Canvas* hält all das strukturiert fest. Damit ein Vorhaben *geordnet* abläuft, helfen *Prozessmodelle* wie *CRISP-DM* und der *End-to-End*-Workflow eines Machine-Learning-Projekts.
Lernziel: Nach Abschluss dieses Teils kann der/die Absolvent:in Data Thinking beschreiben, typische Gründe für das Scheitern von Datenprojekten benennen, ein Datenproblem mithilfe eines Canvas strukturieren sowie die Phasen von CRISP-DM und eines End-to-End-ML-Projekts erklären.
Input:

- Folien *Data Thinking & Canvas* (inkl. Prozessmodelle)
- Video zum Modul

Output:

- Sie können ein Datenproblem klar formulieren, die richtigen Fragen stellen und den Ablauf eines datengetriebenen Projekts (CRISP-DM / End-to-End) einordnen.

Lernerfolg: Bestimmen Sie Ihren Lernerfolg.

Aufgabe 1: Wiederholung

Fassen Sie in eigenen Worten zusammen, was Sie gelernt haben. Welche Elemente gehören zum Data Thinking (Problemdefinition, Beteiligte, Wert, Daten, Strategie)? Welche Phasen umfasst der CRISP-DM-Prozess bzw. ein End-to-End-ML-Projekt, und wozu dient jede Phase?

Antwort

Aufgabe 2: Single Choice

Kreuzen Sie pro Frage die **eine** richtige Antwort an.

1. Welche Phase ist **nicht** Teil des CRISP-DM-Prozesses?

   ☐ Datenaufbereitung

   ☐ Modellierung

   ☐ Software-Entwicklung

   ☐ Evaluierung

2. Was ist das Ziel der Phase „Business Understanding" im CRISP-DM-Modell?

   ☐ Die Daten für Machine-Learning-Algorithmen vorzubereiten.

   ☐ Den geschäftlichen Kontext und die Ziele zu verstehen.

   ☐ Datenmodelle zu evaluieren.

   ☐ Ein endgültiges Modell bereitzustellen.

3. Welcher Schritt ist Teil eines End-to-End Machine Learning Projekts?

   ☐ Daten visualisieren, um Erkenntnisse zu gewinnen.

   ☐ Daten löschen, die nicht benötigt werden.

   ☐ Ein endgültiges Modell auswählen, ohne es zu testen.

   ☐ Die Ergebnisse direkt in einem Bericht dokumentieren.

4. Was ist der letzte Schritt in einem End-to-End Machine Learning Projekt?

   ☐ Die Daten vorverarbeiten.

   ☐ Die Lösung präsentieren.

   ☐ Das System starten, überwachen und warten.

   ☐ Modelle vergleichen und bewerten.

5. Welche Aussage beschreibt die Datenvorbereitung in einem Machine-Learning-Prozess korrekt?

   ☐ Daten werden in Kategorien eingeteilt, um Muster zu erkennen.

   ☐ Daten werden bereinigt und transformiert, um sie für Algorithmen nutzbar zu machen.

   ☐ Daten werden ausschliesslich zur Visualisierung verwendet.

   ☐ Daten werden direkt für die Modellierung verwendet, ohne Anpassungen vorzunehmen.

Aufgabe 3: Offene Fragen

1. Warum scheitern viele Data-Science-Projekte? Erläutern Sie die Aussage „A problem well stated is a problem half solved" (Kettering).

Antwort

2. Was versteht man unter *Data Thinking*, und wie unterscheidet es sich von einem rein technik-/modellgetriebenen Vorgehen?

Antwort

3. Wozu dient ein *Data(-Leadership)-Canvas*? Nennen Sie zentrale Elemente, die hineingehören, und begründen Sie warum.

Antwort

4. Welche Phasen umfasst das CRISP-DM-Modell, und was ist das Ziel jeder Phase?

Antwort

5. Wie unterscheidet sich ein End-to-End Machine Learning Projekt von einem traditionellen Softwareprojekt?

Antwort

6. Beschreiben Sie die Rolle der Modellbewertung in einem Machine-Learning-Projekt.

Antwort

7. Warum ist die Phase „Deployment" entscheidend, und welche Schritte umfasst sie?

Antwort