

Datensätze für Projekt

Datensatz	Beschreibung
Attrition - Train (in Orange3 enthalten)	<p>In diesem Datensatz lassen sich Faktoren, die zur Fluktuation von Arbeitnehmern führen, entdecken und Fragen wie nach dem Einfluss der Entfernung zum Arbeitsort oder dem durchschnittlichen Monatseinkommen untersuchen. Es handelt sich dabei um einen fiktiven Datensatz, der von IBM erstellt wurde. Aufgabe ist es, ein Modell zu trainieren, das abwanderungswillige Mitarbeiter erkennt.</p> <p>Ziel: Abwanderungswillige Mitarbeiter vorhersagen</p>
Bank Marketing (in Orange3 enthalten)	<p>Diese Daten entstammen den Direktmarketing-Kampagnen (Telefonanrufe) einer portugiesischen Bank. Aufgabe ist es, auf Basis dieser Daten vorherzusagen, ob ein Kunde eine Festgeldanlage tätigen möchte. Dazu stehen Kundenprofile zur Verfügung (bspw. mit Alter, Berufstyp oder Informationen über frühere Kredite).</p> <p>Ziel: Potenzielle Kunden für Festgeldanlagen vorhersagen</p>
Breast Cancer (in Orange3 enthalten)	<p>Dieser Datensatz enthält Diagnoseinformationen über Brustkrebs. Aufgabe ist es, auf Basis von (bis zu) 9 Merkmalen vorherzusagen, ob Brustkrebs vorliegt.</p> <p>Ziel: Brustkrebs vorhersagen</p>
Adult (in Orange3 enthalten)	<p>Dieser Datensatz enthält Ergebnisse der amerikanischen Volkszählung von 1994. Aufgabe ist es, ein Modell zu trainieren, das entscheiden kann, ob eine Person mehr als 50 Tausend Dollar pro Jahr verdient.</p> <p>Ziel: Vorhersagen, wer mehr als 50.000\$ pro Jahr verdient</p>
Student Performance Data Set	<p>Dieser Datensatz enthält Leistungen von Schülerinnen und Schüler zweier portugiesischer Schulen. Aufgabe ist es, basierend auf demografischen, sozialen und schulischen Merkmalen vorherzusagen, ob der/die Schüler/in den Kurs bestehen wird</p> <p><i>Dateien finden sich im Begleitmaterial</i></p> <p>Ziel: Bestehen vorhersagen</p>