

➤ Einfaches Identifizieren von Gruppen und Vorhersagen von Gruppenzugehörigkeit

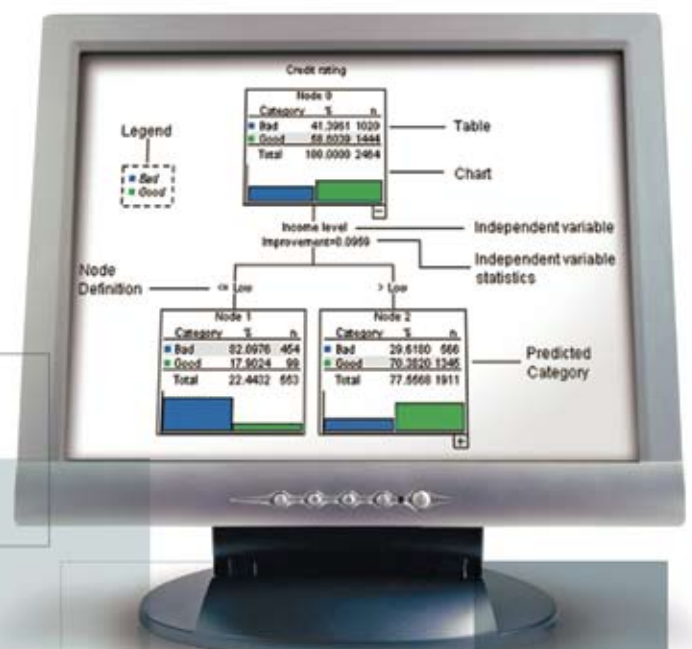
Mit dem Zusatzmodul PASW Classification Trees* können Sie direkt in PASW Statistics Base* Klassifikations- und Entscheidungsbäume erstellen, mit denen Sie leichter Gruppen identifizieren, Beziehungen zwischen Gruppen aufdecken und künftige Ereignisse vorhersagen können. Sie können Klassifizierungs- und Entscheidungsbäume zur Segmentierung, Schichtung, Vorhersage und Dimensionsreduktion sowie zum Variablen-Screening, Ermitteln von Wechselwirkungen, Zusammenführen von Kategorien und Diskretisieren stetiger Variablen verwenden.

Mit übersichtlichen Diagrammen können Sie kategoriale Ergebnisse einem Publikum ohne entsprechende Fachkenntnisse verständlich präsentieren. Diese Baumstrukturen ermöglichen Ihnen, Ihre Ergebnisse zu analysieren und den Modellfluss visuell zu bestimmen. Die übersichtlichen Entscheidungsbäume decken bestimmte Untergruppen und Beziehungen auf, die mit herkömmlichen statistischen Verfahren möglicherweise nicht erkannt werden. Die Daten sind in den Klassifikationsbäumen auf Verzweigungen und Knoten verteilt. So sind Klassengrenzen gleich zu erkennen.

*PASW Decision Trees, PASW Statistics Base und PASW Statistics Base Server, früher bekannt als SPSS Decision Trees, SPSS Statistics Base und SPSS Statistics Base Server sind Teil des SPSS Inc.'s Predictive Analytics Software Portfolios.

In folgenden Szenarien können Sie PASW Decision Trees einsetzen:

- Datenbankgestütztes Marketing
 - Auswählen einer Responsevariable zum Segmentieren des Kundenbestands (z. B. Kunden, die auf eine Testmail geantwortet/nicht geantwortet haben, Kunden mit hohen, mittleren und niedrigen Umsätzen oder Erstkunden, die eine Dienstleistung verlängert haben, und solche, bei denen dies nicht der Fall war)
 - Erstellen von Profilen für Gruppen auf der Grundlage von anderen Attributen, z. B. demografischen Angaben oder Kundenaktivität
 - Abstimmen von neuen Werbeaktionen auf eine bestimmte Untergruppe, Kostenreduktion und Optimierung des Anlageertrags (ROI)



■ **Marktforschung**

- Auswerten von Umfragen zur Zufriedenheit von Kunden, Mitarbeitern oder Neukunden
- Auswählen einer Variablen zum Messen der Zufriedenheit (z. B. auf einer Skale von 1 bis 5)
- Erstellen von Profilen von Zufriedenheitswerten anhand der Antworten auf andere Fragen
- Ändern von Faktoren wie Arbeitsumgebung oder Produktqualität, die Auswirkungen auf die Zufriedenheit haben können

■ **Kreditrisiko-Bewertung**

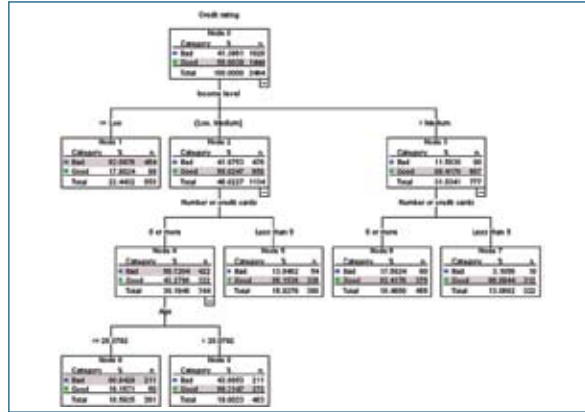
- Bestimmen von Risikogruppen (hoch, mittel oder niedrig)
- Erstellen von Profilen von Risikogruppen auf der Grundlage von Kundeninformationen, z. B. Kontobewegungen
- Anbieten der richtigen Kreditlinie für die richtigen Bewerber auf Grundlage von Risikogruppen

■ **Bewertung von Förderprogrammen**

- Auswählen einer Variablen mit gewünschtem und nicht gewünschtem Ergebnis (z. B. erfolgreicher Abschluss einer Arbeitsbeschaffungsmaßnahme)
- Aufdecken erfolgsrelevanter Faktoren - auf der Grundlage von Bewerberinformationen
- Anpassen von neuen Programmen für ein breiteres Publikum

■ **Marketing im öffentlichen Sektor**

- Auswählen einer Responsevariablen zum Segmentieren des Kundenbestands (z. B. potenzielle Universitäts-Bewerber, die sich wirklich beworben haben, im Gegensatz zu solchen, bei denen dies nicht der Fall ist)
- Erstellen von Profilen für Gruppen auf der Grundlage von anderen Attributen, z. B. demografischen Angaben oder Kundenaktivität
- Kostenreduktion und Optimierung des ROI durch Abstimmen von neuen Werbeaktionen auf eine bestimmte Untergruppe

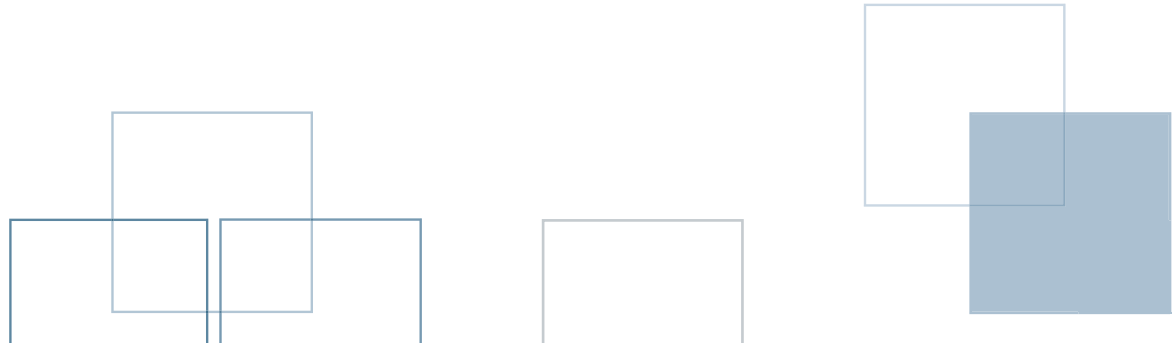


Decken Sie mit anschaulichen Bäumen Beziehungen auf, die in Ihren Daten zurzeit nicht erkennbar sind. Die Diagramme, Tabellen und Grafiken in PASW Decision Trees können problemlos interpretiert werden.

	age	gender	marital	ed	marital	PredictedValue
1	45 m		4.00	3	1	83.16
2	42 m		2.00	3	0	17.86
3	39 m		1.00	4	0	10.26
4	35 f		2.00	4	1	17.26
5	42 f		4.00	3	0	80.16
6	36 m		3.00	3	0	30.21
7	29 m		2.00	3	1	16.26
8	39 m		1.00	4	1	10.26
9	39 m		2.00	3	1	16.66
10	49 f		4.00	2	0	84.90
11	34 f		1.00	3	0	10.26
12	47 m		1.00	3	1	10.26
13	58 m		3.00	4	0	30.21

Verwenden Sie die resultierenden Baummodelle, um direkt in PASW Statistics Fälle zu bewerten.

PASW Decision Trees ist in einer Ausführung als reiner Client verfügbar. Zur Steigerung der Leistung und Skalierbarkeit kann es jedoch auch in einer Client-/Server-Installation mit PASW Statistics Base Server* verwendet werden.



Auswahl zwischen vier Algorithmen für Entscheidungs­bäume

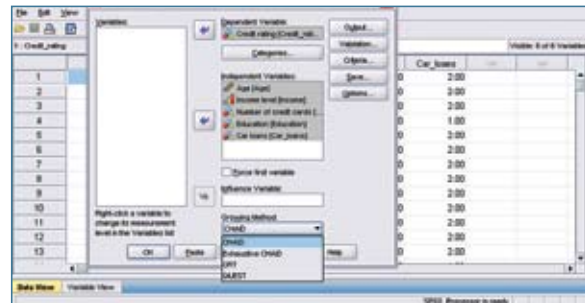
PASW Decision Trees umfasst vier bewährte Algorithmen für den Aufbau des Baums:

- **CHAID** — Ein schnelles, Statistik-basiertes Entscheidungsbaumverfahren, das Mehrfach-Verzweigungen unterstützt und jeweils Segmente und Profile für Gruppen erstellt.
- **Exhaustive CHAID** — Eine Modifikation von CHAID, mit der alle möglichen Teilungen für jeden Prädiktor untersucht werden.
- **Classification & Regression Trees (C&RT)** — Ein vollständiger binärer Entscheidungsbaumalgorithmus, mit dem Daten partitioniert und genaue homogene Untergruppen erzeugt werden.
- **QUEST** — Ein statistischer Algorithmus, mit dem Variablen ohne Verzerrung ausgewählt und schnell und effizient genaue binäre Entscheidungs­bäume erstellt werden.

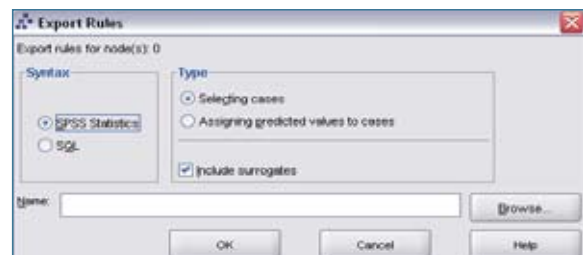
Mit vier Algorithmen haben Sie die Möglichkeit, unterschiedliche Methoden für den Aufbau des Baums auszu­probieren und das Verfahren zu ermitteln, das am besten für die Daten geeignet ist.

Erweiterte Ergebnisse mit weitergehenden Analysen in PASW Statistics

PASW Decision Trees wird über die PASW Statistics Base-Benutzer­oberfläche verwendet. Dadurch können Sie bequem Klassifikations­bäume erstellen und anhand der Ergebnisse Fälle unmittelbar in den Daten segmentieren und gruppieren. Sie müssen nicht zwischen PASW Statistics Base und anderer Software wechseln. Darüber hinaus können Sie mit der PASW Statistics-Syntax, mit SQL-Anweisungen oder einfachem Text (über die Syntax) Regeln für die Auswahl oder die Klassifikation/Vorhersage



Erstellen Sie Baummodelle in PASW Statistics mit „CHAID“, „Exhaustive CHAID“, „C&RT“ oder „QUEST“.



Wählen Sie direkt in den Daten aus den Modellergebnissen Fälle aus, oder weisen Sie Vorhersagen zu, oder exportieren Sie Regeln für die spätere Verwendung.

generieren. Sie können diese Regeln im Viewer anzeigen und sie für die spätere Verwendung in einer externen Datei speichern, um Vorhersagen über einzelne und neue Fälle zu treffen. Wenn Sie anhand der Ergebnisse andere Datendateien bewerten möchten, können Sie Informationen direkt aus dem Baummodell in die Daten schreiben oder XML-Modelle erstellen, die in PASW Statistics Base Server verwendet werden können.

Funktionen

Bäume

- Anzeigen von Baumdiagrammen, Baumstrukturen, Balkendiagrammen und Datentabellen
- Bequemes Erstellen von Bäumen. Die benutzerfreundliche Oberfläche bietet folgende Optionen:
 - Messniveau (nominal, ordinal und stetig)
 - Unabhängige Variablen
 - Abhängige Variablen
 - Einflussvariablen
 - Aufbaumethode
 - Ausgabeeinrichtung, u. a. Bäume, Statistiken, Diagramme und Regeln
 - Stichproben- oder Kreuzvalidierung
 - Kriterien für das Anhalten
 - Gespeicherte Variablen, u. a. vorhergesagte Werte für Wahrscheinlichkeiten
- Auswählen aus vier Methoden für den Aufbau des Baums
- Anzeigen von Knoten mit verschiedenen Methoden: Anzeigen von Balkendiagrammen und/oder Tabellen der Zielvariablen in jedem Knoten
- Erweitern und Reduzieren von Verzweigungen und Ändern sonstiger Darstellungseigenschaften, beispielsweise von Schriftarten und Farben
- Anzeigen und Ausdrucken von Bäumen
- Festlegen der exakten Vergrößerungsstufe (in Prozent) für die Anzeige von visuellen Baummodellen in der Benutzeroberfläche
- Automatisieren der Erstellung von Bäumen im Produktionsmodus
 - Automatisches Generieren von Syntax aus der Oberfläche
- Erzwingen der Aufnahme eines Prädiktors in das Modell
- Eingeben von A-priori-Wahrscheinlichkeiten, Fehlklassifikationskosten, Erträgen, Ausgaben und Skalenwerten

Algorithmen für den Aufbau des Baums

- Durchführen von Analysen mit einem von vier leistungsfähigen Algorithmen für den Aufbau des Baums:
 - CHAID-Methode nach Kass (1980)
 - Exhaustive CHAID-Methode nach Biggs, de Ville und Suen (1991)
 - Classification & Regression Trees (C&RT) nach Breiman, Friedman, Olshen und Stone (1984)
 - QUEST-Methode nach Loh und Shih (1997)
- Behandeln von fehlenden Prädiktorwerten mit einer der zwei folgenden Methoden: Zuweisen zu einer Kategorie oder Vorschreiben unter Verwendung einer Ersatzvariablen
- Diskretisieren stetiger Prädiktorvariablen gemäß der angegebenen Kategorieanzahl
- Beschneidungsfunktionen für „C&RT“ und „QUEST“
- Zufallsstichproben aus Quelldaten für Stichprobenvalidierung oder Verwendung einer Variablen zum Teilen der Stichprobe

Modellauswertung

- Generieren von Risiko- und Klassifikationstabellen
- Zusammenfassen der Knotenleistung mit Auswertungsdiagrammen und -tabellen, mit deren Hilfe die besten Segmente identifiziert werden können
 - Gewinne
 - Index (Gain)
 - Response
 - Mittelwert
 - Durchschnittlicher Profit
 - ROI
- Aufteilen von Daten in Trainings- und Testdaten zum Überprüfen der Genauigkeit
- Anzeigen von Auswertungsdiagrammen bzw. Klassifikationsregeln für ausgewählte Knoten im Knotenübersichtsfenster

Bereitstellung

- Export:
 - Baumdiagramme, Diagramme und Tabellen. Folgende Exportformate sind verfügbar: HTML-, Text-, Word/RTF- und Excel®-Dateien sowie PDF
- Speichern von Informationen aus dem Modell als Variablen in der Arbeitsdatei
- Exportieren von Entscheidungsregeln zum Definieren ausgewählter Segmente in SQL zum Bewerten von Datenbanken, als PASW Statistics-Syntax zum Bewerten von PASW Statistics-Dateien oder als einfachen Text (über Syntax)
- Exportieren von Bäumen als XML-Modelle, um mit PASW Statistics Base Server und SmartScore® neue Fälle oder Datendateien zu bewerten
- Veröffentlichen von Bäumen als Bilder und Veröffentlichen von Tabellen als statische oder interaktive Tabellen in SmartViewer® Web Server™
- Auswählen bestimmter Segmente in der Arbeitsdatei über Baumknoten und Ausführen weiterer Analysen für zusätzliche Erkenntnisse

Systemanforderungen

- Software: PASW Statistics Base* 17.0
- Andere Systemanforderungen können je nach Plattform abweichen.

Die endgültige Version kann geänderte Funktionen enthalten.



Weitere Informationen erhalten Sie auf unserer Website www.spss.de. SPSS GmbH Software – Theresienhöhe 13 – 80339 München – Tel. +49.89 48 90 74-0, Fax +49.89.448 31 15.

SPSS ist eine eingetragene Marke, und alle weiteren genannten SPSS Inc.-Produkte sind Marken von SPSS Inc. Alle anderen Namen sind Marken ihrer jeweiligen Eigentümer. © 2009 SPSS Inc. Alle Rechte vorbehalten. SDF1702SPC-0209-DE