



ハイライト

IBM SPSS Decision Trees を使用して、以下を行えます。

- 分類ツリーを使用して、高度に視覚化された方法で、グループ、セグメント、およびパターンを識別
 - CHAID、Exhaustive CHAID、C&RT、および QUEST の中から、データに最適なモデルを選択
 - 専門知識がない人にも分かりやすく、直観的に理解できる方法で結果を表示
 - ツリーから得られる情報 (ターミナル・ノード番号、予測値、予測確率など) をデータの新しい変数として保存
-

IBM SPSS Decision Trees

グループの特定と結果の予測が容易に

IBM SPSS Decision Trees は、分類 (classification) ツリーおよびデシジョン・ツリー (決定木) を作成して、グループの識別の向上、グループ間の関係の発見、将来の事象の予測を支援します。

分類ツリーとデシジョン・ツリーは、以下に利用できます。

- セグメント化
- 層化
- 予測
- データの縮約および変数のスクリーニング
- 交互作用の特定
- カテゴリー結合
- 連続変数の離散化

高度に視覚化されたツリーに、分類別の結果を直観的に理解できる方法で表示するため、専門知識のない人に対しても分類別の結果を分かりやすく説明できます。これらのツリーによって、結果を検討したり、モデルの流れを視覚的に把握することができます。結果を視覚的に表現することで、従来の統計では発見できなかったサブグループや関係性を検出できます。分類ツリーはデータを枝や葉 (ノード) に分岐させるため、グループがどこで分岐/終了するのかを容易に把握できます。

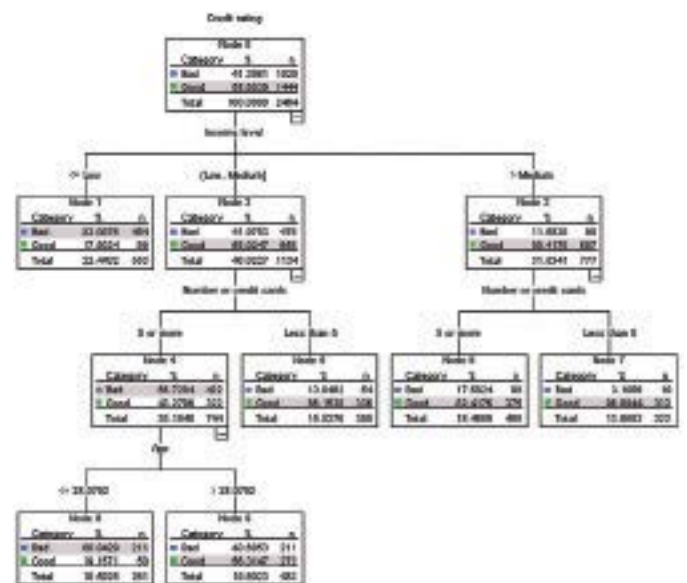


SPSS Decision Trees は、以下のようなさまざまな分野で利用できます。

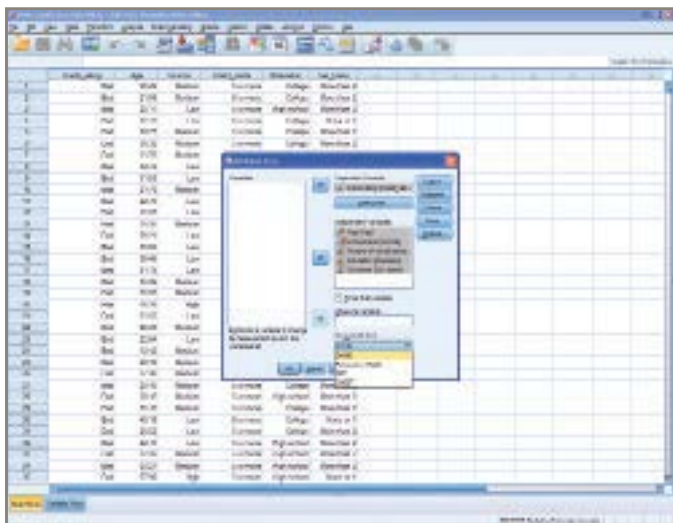
- データベース・マーケティング:
 - 顧客ベースでセグメント化するための応答変数 (試験的 DM に対する反応者/非反応者、高/中/低収益顧客、サービスを延長した/延長しなかった新規顧客) を 1 つ選択
 - 人口統計量や顧客行動などの他の属性に基づいて、グループをプロファイル化
 - 1 つのサブグループに特化した新しい販促活動を作成し、経費削減および投資収益率 (ROI) の向上を目指す
- マーケット・リサーチ:
 - 顧客、従業員、またはその他のタイプの満足度調査の実施
 - 満足度を測定する変数 (5 段階評価など) を選択
 - 他の質問に対する回答に基づいて、満足度のレベルをプロファイル化
 - 要因 (労働環境、製品品質など) を変更し、これが満足度にどのように影響するかを把握
- 信用リスクのスコアリング:
 - 信用リスク・グループ (高、中、低) の特定
 - 口座状況等の顧客情報に基づいて、信用リスク・グループをプロファイル化
 - 信用リスク・グループに基づいて、適切な申込者に対して適切な与信を行うことが可能
- プログラムの対象選定:
 - 望ましい/望ましくない結果 (職業訓練課程の修了等) を表す変数を 1 つ選択
 - 申込者の情報に基づき、成功をもたらす要因を明らかにする
 - より多くの人の必要を満たすために新規プログラムをカスタマイズ

- 公共機関のマーケティング:
 - 顧客ベースをセグメント化するための応答変数 (例えば、大学入学申込者のうち、実際に入学した人と入学しなかった人を表す変数など) を 1 つ選択
 - 人口統計量や顧客行動などの他の属性に基づいて、グループをプロファイル化
 - 1 つのサブグループに特化した新しい販促活動を作成し、経費削減および ROI の向上を目指す

SPSS Decision Trees は、クライアント専用ソフトウェアとしてインストールできますが、より高度なパフォーマンスと拡張性が必要な場合は、サーバー版をご利用いただけます。



視覚的に分かりやすいツリーを使用して、これまでデータ内で隠されていた関連性を発見します。SPSS Decision Trees の図、表、およびグラフは、理解しやすいものとなっています。



ツリー・モデルの結果を使用して、直接 IBM SPSS Statistics 内でケースのスコアリングを行うことができます。

4 つのデジジョン・ツリー・アルゴリズムから選択

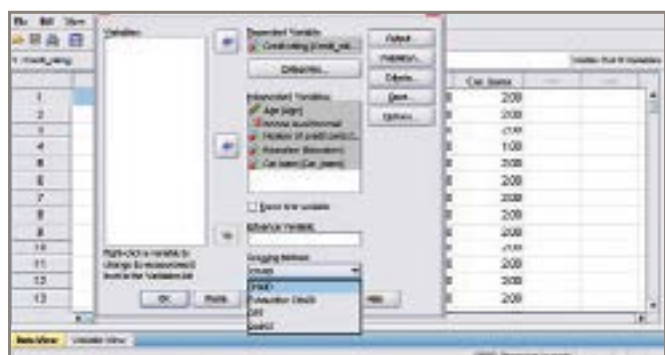
SPSS Decision Trees には、広く使用されているツリー成長アルゴリズムが 4 つ含まれています。

- **CHAID** – 高速な統計的多分木 (マルチウェー・ツリー) アルゴリズムで、データを素早くかつ効率的に探索し、望まれる結果に関してセグメント化ならびにプロファイル化を行う
- **Exhaustive CHAID** – CHAID の修正版で、各予測変数における可能なすべての分岐を検討
- **Classification & regression trees (C&RT)** – 完全な二進木 (バイナリー・ツリー) アルゴリズムで、データを分割し、同質のサブセットを正確に作成
- **QUEST** – 偏り (バイアス) なく変数を選択する統計アルゴリズムで、二進木を高速かつ効率的に作成

これらの 4 つのアルゴリズムを使用して、異なるツリー成長手法を試し、その中から当該データに最適なものを検出できます。

より高い価値を創出するコラボレーション

IBM SPSSソフトウェアを IBM SPSS Collaboration and Deployment Services と統合して使用することで、分析資産の共有および再利用を効率的に行い、社内外のコンプライアンス要件を満たす方法で保護し、分析結果を公開して、より多くのビジネス・ユーザーが閲覧および利用できるようにすることが可能です。



CHAID, Exhaustive CHAID, C&RT, または QUEST を使用して、SPSS Statistics モデル内でツリー・モデルを作成します。

得られた結果を IBM SPSS Statistics Base でさらに分析

SPSS Decision Trees を IBM SPSS Statistics Base と組み合わせると、分類ツリーを容易に作成し、この結果をデータ内で直接使用して、ケースのセグメント化やグループ化を行うことができます。また、選択方法や分類/予測ルールを、SPSS Statistics シンタックス、SQL ステートメント、シンプル・テキスト (シンタックスを使用) の形式で生成できます。このルールをビューアーで表示したり、または外部ファイルに保存して、後で個々の新しいケースを予測する際に使用できます。結果を利用して他のデータ・ファイルのスコアリングを行う場合、ツリー・モデルからデータに直接情報を書き込むか、または IBM SPSS Statistics Base Server で使用できるように XML モデルを作成することができます。

モデルの結果を利用してデータ内でケースを直接選択したり、予測を割り当てたり、あるいはルールをエクスポートして後で使用することができます。

機能

ツリー

- ツリー図、ツリーマップ、棒グラフ、データ表の表示
- 包括的なインターフェースを使用することで、ツリーを簡単に作成。以下の設定が可能:
 - 測定尺度 (名義/順序/連続)
 - 独立変数
 - 従属変数
 - 影響度変数
 - 成長手法
 - 出力設定 (ツリー、統計、グラフ、ルール)
 - 分割サンプル検証または交差検証
 - 停止基準
 - 変数の保存 (予測値、確率、XML モデル)
- 4 つのツリー成長手法から選択
- 複数の方法のいずれかでノードを表示: 目的変数の棒グラフ、表 (いずれかあるいは両方) を各ノード内に表示
- 枝の縮小/展開が可能。フォントや色などの外観上の特性を変更可能
- ツリーの表示と出力
- インターフェース内でツリー・モデルを表示する際に、表示の大きさをパーセントで厳密に設定
- バッチ・モードによる、ツリー作成の自動化
- インターフェースからシンタックスを自動生成
- モデルに使用する予測変数を 1 つ強制的に選択
- 事前確率、誤分類コスト、収益、費用、スケール・スコアを指定

ツリー成長アルゴリズム

- 強力な 4 つのツリー成長アルゴリズムのいずれかを使用した分析の実行:
 - CHAID : Kass による (1980)
 - Exhaustive CHAID : Biggs, de Ville, and Suen による (1991)
 - Classification & Regression Trees (C&RT) : Breiman, Friedman, Olshen, および Stone による (1984)
 - QUEST : Loh および Shih による (1997)
- 予測変数の欠損データを 2 つの方法のいずれかで処理: カテゴリーへの割り当てと代理変数による計算
- 指定されたカテゴリーの個数に基づいて、連続する予測変数を離散化
- C&RT および QUEST の剪定機能
- ソース・データをランダムにサンプリングし、分割サンプル検証や、変数によるサンプルの分割を行う

モデルの評価

- リスク表と分類表を生成
- 評価グラフや表により、ノードのパフォーマンスを要約し、最良のセグメントを特定:
 - ゲイン
 - インデックス (リフト)
 - 応答
 - 平均
 - 平均プロフィット
 - ROI
- データを学習用と検証用に分け、精度を確認
- ノードの要約ウィンドウを使用し、選択されたノードの要約グラフや分類ルールを表示

展開

- エクスポート:
 - ツリー図、グラフ、および表は、HTML、テキスト、Microsoft® Word、Microsoft Excel、RTF、および PDF の形式でエクスポート可能
- モデルからの情報を変数として作業中のデータ・ファイルに保存
- 選択されたセグメントを定義するデジジョン・ルールを SQL としてエクスポートして、データベースのスコアリングを行う。または、IBM SPSS Statistics シンタックスとしてエクスポートして IBM SPSS Statistics ファイルのスコアリングを行う。あるいは、(シンタックスを使用して) シンプル・テキストとしてエクスポートする
- ツリーを XML モデルとしてエクスポートし、IBM SPSS Statistics Server や SPSS Statistics Portal で使用
- ツリーをイメージとして、また表を静的な表またはインタラクティブな表として SPSS Statistics Portal に公開
- より詳細な分析を行うには、ツリー・ノードから作業中のデータ・ファイル内の対象セグメントを選択し、これに対してさらに分析を行う

システム要件

要件はプラットフォームによって異なります。

IBM ビジネス・アナリティクスについて

IBM Business Analytics ソフトウェアは、業績改善に取り組む意思決定者に対し、実践的な洞察を提供します。IBM は、ビジネス・インテリジェンス、予測分析と高度な分析、財務パフォーマンスと戦略の管理、ガバナンス、リスクおよびコンプライアンス (GRC)、そしてアナリティック・アプリケーションからなる包括的なポートフォリオを用意しています。

IBM ソフトウェアは、ビジネスの傾向やパターンあるいは異常の発見、仮定に基づくシナリオの比較、潜在的な脅威や機会の予測、重要なビジネス・リスクの特定および管理、さらには経営資源に関する計画、予算および予測を実現します。IBM の世界中のお客様は、この充実したアナリティクスを使うことで、業績への理解を深める一方、成果への予測を高め、目標への確かな道筋をつけることができます。



日本アイ・ビー・エム株式会社
〒103-8510
東京都中央区日本橋箱崎町 19 番 21 号

IBM のホーム・ページはこちらからご覧になれます。
ibm.com

IBM、IBM ロゴ、ibm.com および SPSS は、世界の多くの国で登録された International Business Machines Corporation の商標です。このような商標は、その他の国においても登録商標またはコモン・ロー上の商標である可能性があります。現時点での IBM の商標リストについては、ibm.com/legal/copytrade.shtml をご覧ください。

Microsoft、Windows、Windows NT および Windows ロゴは、Microsoft Corporation の米国およびその他の国における商標です。

他の会社名、製品名およびサービス名等はそれぞれ各社の商標です。

P26470

© Copyright IBM Corporation 2012
All Rights Reserved.



リサイクル可能

Business Analytics software