

IBM® Analytical Decision Management
バージョン 8 リリース 0

IBM® SPSS® Modeler
Advantage

IBM

お願い

本書および本書で紹介する製品をご使用になる前に、95 ページの『特記事項』に記載されている情報をお読みください。

本書は、IBM Analytical Decision Management バージョン 8 リリース 0 モディフィケーション 0 および新しい版で明記されない限り、以降のすべてのリリースおよびモディフィケーションに適用されます。

お客様の環境によっては、資料中の円記号がバックスラッシュと表示されたり、バックスラッシュが円記号と表示されたりする場合があります。

原典： IBM® Analytical Decision Management
Version 8 Release 0
IBM® SPSS® Modeler Advantage

発行： 日本アイ・ビー・エム株式会社

担当： トランスレーション・サービス・センター

第1刷 2013.10

© Copyright IBM Corporation 2010, 2013.

目次

前書き	v
-----	---

第 1 章 IBM Analytical Decision

Management について 1

製品概要	1
バージョン 8 の新機能	1
前バージョン	3

第 2 章 顧客応答のモデリングの例 9

IBM SPSS Modeler Advantage	9
ダイレクト・メール キャンペーンの場合の利益の最 大化	9
サンプル・ファイル	9
アプリケーションの起動	11
データ・ソースの設定	11
モデルの構築	13
モデルの評価	13
予測ルール・モデル	15
モデルのテスト	15
モデルのスコアリング	16
クラスター・モデリングの例	16
アソシエーション・モデリングの例	18

第 3 章 作業領域のナビゲート 21

表示言語の変更	21
アプリケーションの起動	21
ホーム・ページ	22
ギャラリー	23
設定の指定	23
プロジェクト、モデル、およびルールのオープンと 保存	25
IBM SPSS Collaboration and Deployment Services Repository への保存	25
ラベルの操作	26
プロジェクトのロック	27
プロパティ	27

第 4 章 データ・ソースの管理 29

データのプレビュー	31
データ概要	32
測定の尺度	32
データ・フィールド値の編集	33
新規データ・ソースの作成	34
ファイル・ソース	34
データベース・ソース	37
「Enterprise View」ソース	38
データ・ビュー・ソース	38
IBM Cognos BI ソース	39
Analytic Server ソース	41
追加フィールドの追加	42

入力フィールドの選択	42
式エディター	43
フィールドのマップ	43

第 5 章 ルールの操作 45

ルールの作成	45
セグメント・ルールの定義	45
選択ルールの定義	47
ルールのエクスポートおよび再使用	48
注釈の追加	50
IBM Operational Decision Management からのルール の使用	50

第 6 章 予測モデルの作成 53

予測モデルについて	53
モデルの作成	53
予測モデルの作成	53
予測モデル作成の結果	54
予測ルール・モデルの作成	55
クラスターリング・モデルの作成	57
アソシエーション・モデルの作成	58
オプションのモデル設定	60
モデルの評価	63
最大利益のシミュレート	65
モデルのテスト	66
アプリケーションでのモデル・スコアの使用	66

第 7 章 スコアリング・モデル 69

データベース・テーブル、ファイル、Analytic Server、または、IBM Cognos BI サーバーのモデル のスコアリング	69
データおよびサブセットの選択	69
スコアリングするレコードの選択	70
スコアリングの宛先の選択	70
出力フィールドの選択	74
モデルのスコアリング	74
スコアリング要約レポート	75

第 8 章 アプリケーションの管理 77

一般オプション	78
アプリケーション・ショートカットの制御	79
ホーム・ページの非表示設定	79
データ・オプションのロック	79
スコアリング宛先オプションのロック	80
外部ルールの作成に使用するプロジェクト・メタデ ータのダウンロード	80
XML スキーマ定義の編集	82

第 9 章 ヘルプの入手 89

利用可能な資料	89
---------	----

付録 A. ストリームの共有	91	索引	97
付録 B. アクセシビリティ	93		
特記事項	95		
商標	96		

前書き

IBM Business Analytics について

IBM Business Analytics ソフトウェアは、意思決定者が業績向上を図る際に頼りになる、完全で一貫性のある正確な情報を提供します。ビジネス・インテリジェンス、予測分析、財務業績/戦略管理、および分析アプリケーションからなる包括的なポートフォリオにより、現在の業績に関する明確で直接的かつ実用的な洞察を得て、将来の成果を予測できるようになります。豊富な業界ソリューション、実証済みの手法、および専門のサービスを組み合わせることにより、あらゆる規模の企業が、高い生産性を実現するとともに、自信を持って意思決定を自動化し、優れた結果を出すことができます。

このポートフォリオの一部である IBM SPSS Predictive Analytics ソフトウェアにより、企業は将来のイベントを予測し、その洞察に基づいて事前対処的に行動して、ビジネス成果を高めることができます。全世界の民間、行政、学術の各分野のお客様が、IBM® SPSS® テクノロジーを活用して、不正の削減とリスクの軽減を図りながら、顧客の獲得、維持、拡大での競争優位性を高めています。IBM SPSS ソフトウェアを日々の業務に取り入れることで、企業は予測的に対応できるようになり、ビジネス目標を達成し、高い競争優位性を実現するように意思決定を方向付け、自動化することが可能になります。詳細情報、および担当員の連絡先については、<http://www.ibm.com/spss> をご覧ください。

テクニカル・サポート

保守契約があるお客様は、テクニカル・サポートをご利用いただけます。IBM Corp. 製品の使用方法、または対応するハードウェア環境へのインストールについてサポートが必要な場合は、テクニカル・サポートにご連絡ください。テクニカル・サポートをご利用になる場合は、IBM Corp. Web サイト (<http://www.ibm.com/support>) をご覧ください。ご利用の際には、ご本人、所属する組織、およびサポート契約を確認できるものをご用意ください。

第 1 章 IBM Analytical Decision Management について

製品概要

IBM Analytical Decision Management では、予測分析の利点を実際のビジネス上の課題に活かして、顧客や業界に合わせたカスタム・アプリケーションを作成することができます。一般にアプリケーションは特定の課題を解決するように構成されますが、ベースとなる機能はすべてのアプリケーションで共通です。

- ビジネス・ルールを使用して自動的に意思決定を行います。
- 予測モデルを使用して洞察を加えます。
- 優先順位、最適化、またはシミュレーションを使用して、上記の機能を基に最良の決定を導きます。

特定のビジネス上の課題を解決するために調整された、多数のパッケージ・アプリケーションが用意されています。詳しくは、営業担当員にお問い合わせください。

バージョン 8 の新機能

IBM Analytical Decision Management 8 (8.0.0.0) は、日常のビジネス・プロセスに分析を展開するユーザーの能力を拡張し、より広範囲の意思決定をかつてないほど優れた洞察と制御によりサポートします。

- 高度なシミュレーション機能を使用して、ビジネス成果に関するルール、モデル、および最適化の複合的な影響を微調整および視覚化します。
- 事前にビルドされた新規アプリケーション、および価格最適化や分類最適化などの新規最適化手法により、より広範囲の問題を解決します。
- IBM Operational Decision Management との改善された統合、および IBM SPSS Analytic Server 統合によるビッグ・データのサポートを提供します。

以下は主要な新機能についての説明です。

新規インストール・テクノロジー。 製品のインストールが簡素化され、改善されました。新たに、Installation Manager テクノロジーが使用されるようになりました。詳しくは、「*IBM Analytical Decision Management* インストール・ガイド」を参照してください。

ユーザー・インターフェースの機能拡張。 ユーザー・インターフェースに多くの変更が行われ、使いやすさが増しました。

IBM Analytical Decision Management for Predictive Maintenance. IBM Analytical Decision Management for Predictive Maintenance は、設備の障害が発生する前にその障害を予測して防止するための最適化を使用する、事前ビルドされた新規アプリケーションです。ユーザーは、障害を予測し、適切な修正アクションを適時にスケジュールすることで、コストのかかるダウン時間を最小限に抑え、希少なインフラストラクチャーおよびリソースを最大限に活用できます。ユーザーは、どの部分で最初に障害が発生しそうかを予測し、その部分を適切なタイミングで置き換えることが可能となります。マシンでの障害発生時期を予測することは、利益を最大限確保し、マシンの障害および稼働停止による損失を回避することに役立ちます。

IBM Analytical Decision Management for Demand Optimization. IBM Analytical Decision Management for Demand Optimization は、各ストアに各製品の適正量を割り振ることで、コストのかかる過剰在庫および在庫不足を回避するための最適化を使用する、事前ビルドされた新規アプリケーションです。各販路に適正量の在庫を割り振ることで、利益を最大化し、在庫不足や補充費用による損失を回避できます。

シミュレーションの機能拡張。新規シミュレーション機能は、以前の「シミュレーション」ボタンの機能 (以前の「定義」タブ上) と以前の「WhatIf」機能 (以前の「最適化」、「優先順位付け」、または「結合」タブ上) を 1 つの機能に結合します。新規および改善されたシミュレーション機能は、ユーザー・インターフェースの右上隅に独自のツールバーを備えており、各シナリオをプロジェクトと共に保存して、セッションをまたがって持続できるようにします。

展開とスコアリング。「展開」および「スコア」タブにさまざまな機能拡張が行われました。

アプリケーション URL。IBM SPSS Collaboration and Deployment Services では、IBM Analytical Decision Management を起動するための URL を含む、Web ベース・アプリケーションの URL をカスタマイズする機能が追加されました。詳しくは管理者にお問い合わせください。詳しい情報については、*IBM SPSS Collaboration and Deployment Services* リポジトリのインストールおよび設定ガイドの「アプリケーションのコンテキスト・ルート」の章を参照してください。

Statistics ファイルの拡張。IBM SPSS Statistics のパスワード保護された圧縮データ・ファイル (.sav および .zsav) が IBM Analytical Decision Management でサポートされるようになりました。

ツリー・ビューアーの拡張 IBM SPSS Modeler Advantage で個々のモデル結果を表示すると、各モデルの名前がリンクになっており、このリンクからモデルに関する情報 (ターゲット・フィールドや入力フィールドなど) を含む新規ダイアログを開くことができます。選択したモデルがツリー・モデルの場合は、ツリー・ビューアーも表示されます。ツリー・ビューアーは機能拡張されています。ツリーのズームインおよびズームアウト、ツリーの方向の変更、大規模ツリーのナビゲート、およびツリー内での個々のノードの展開または省略などのアクションを実行できるようになりました。

「最適化」タブの「詳細」オプション。IBM Analytical Decision Management for Campaign Optimization、IBM Analytical Decision Management for Predictive Maintenance、および IBM Analytical Decision Management for Demand Optimization など、Complex Mathematical Optimization (CPLEX) を使用するアプリケーションでは、CPLEX 最適化アルゴリズムに関連する詳細オプションが使用可能になりました。最適化プログラミング言語 (OPL) モデル・ファイルのエクスポート、およびカスタム設定の追加が可能になりました。

データ・ビュー・ソース。IBM SPSS Collaboration and Deployment Services 分析データ・ビューで定義されたデータをストリームに組み込むための、新規データ・ビュー・ノードが IBM SPSS Modeler 16 に追加されました。この新しいデータ・ビュー・データ・ソースは、IBM Analytical Decision Management の「データ」タブでも使用できます。IBM Operational Decision Management ルールで使用するための接続を作成できます。階層データ・モデルを使用する複雑なルールについては、パフォーマンス低下を招く恐れがあるため管理者がそれらのルールの使用を制限している場合があります。

V 8 より前のバージョンの外部ルールでは、IBM Analytical Decision Management プロジェクト・データ・モデルと一致するフラット入力パラメーターが必要でした。しかし、この新規の分析データ・ビュー・サポートにより、外部ルールで階層入力パラメーターを使用できるようになりました。詳しくは、トピック 38 ページの『データ・ビュー・ソース』および 79 ページの『データ・オプションのロック』、および「アプリケーション・デザイナーズ・ガイド」の第 5 章 を参照してください。

IBM SPSS Analytic Server 統合。IBM SPSS Modeler 16 では、IBM SPSS Analytic Server が提供され、Hadoop Distributed File System (HDFS) をデータ・ソースとして使用できるようになります。IBM SPSS Modeler ユーザーは、HDFS Analytic Server ソース・ノードを使用してストリームを作成して Hadoop ビッグ・データ・システムに基づいてモデルを構築できます。IBM Analytical Decision Management ユーザーは、これらのモデル・ストリームを IBM SPSS Modeler Advantage で開いて表示できます (Analytic Server データ・ソースを使用して新規プロジェクトを作成することはできません)。

「アプリケーション・デザイナーズ・ガイド」の新しい章。「IBM Analytical Decision Management アプリケーション・デザイナーズ・ガイド」に新しい章が追加されました。この新しい章には、カスタムの最適化アプリケーションの構成例が記載されています。標準の IBM Analytical Decision Management for Campaign Optimization アプリケーションでは、*CampaignOptimization.mod* という名前の事前構成された最適化プログラミング言語 (OPL) モデル・ファイルを使用します。場合によっては、標準の「キャンペーンの最適化」アプリケーションとその OPL モデルが、目の前にあるビジネス上の課題の目的に適していないことがあります。この新しい章では、架空の企業について、IBM ILOG CPLEX Optimization Studio を使用して作成されたカスタム OPL モデルの例、およびカスタムの IBM Analytical Decision Management 最適化アプリケーションの例を紹介します。

表形式の評価結果。 評価の実行時に、新規の「表形式の結果を含める (Include tabular results)」オプションが使用可能になりました。表形式の結果を含む評価の実行にはより多くの時間がかかるため、このオプションを選択解除するとパフォーマンスが向上します。

名義型フィールドの最大数の新規設定。 「名義型」データ・フィールドに指定できるメンバーの最大数を設定するための新規設定が「プロパティ」ダイアログに追加されました。これは「データ」タブで使用されます。詳しくは、27 ページの『プロパティ』を参照してください。

シングル・サインオン。 ODBC 接続で、IBM SPSS Modeler Server、および IBM Analytical Decision Management との ODBC データベース接続のシングル・サインオン (SSO) が使用できるようになりました。いくつかの構成ステップが必要になります。インストール資料の『セキュリティの考慮事項』のトピックを参照してください。

IBM Analytical Decision Management 7 Fix Pack 1 (7.0.0.1) にも新機能が追加されていることに注意してください。詳細については、『バージョン 7.0.0.1』を参照してください。

前バージョン

バージョン 7.0.0.1

IBM Analytical Decision Management 7 Fix Pack 1 (7.0.0.1) には、以下の機能が追加されました。

データ概要。 すべてのアプリケーションの「データ」タブ、IBM SPSS Modeler Advantage の「モデルの作成」タブと「スコア」タブ、および「ただちにスコアリング」機能が含まれているアプリケーションの「展開」タブで、新しいデータ概要機能を使用できます。詳しくは、32 ページの『データ概要』のトピックを参照してください。この機能は、IBM SPSS Modeler Server 15 フィックスパック 1 (15.0.0.1) でのみ使用できます。バージョン 15 では使用できません。

スコアリング要約レポート。 プロジェクトのスコアリングを行った後、「スコアリング要約レポート」セクションに、データのスコアリング先、ファイル・サイズ (該当する場合)、データベースおよびテーブル名 (該当する場合)、モデル名またはプロジェクト名、スコアリングされたレコード数、および場合によってはモデルをまとめた図および/またはテーブルなどの詳細が示されます。「スコアリング要約レポート」セクションは、IBM SPSS Modeler Advantage で表示されるほか、「ただちにスコアリング」機能を備えたアプリケーションの「展開」タブに表示されます。ファイル・サイズは、IBM SPSS Modeler Server 15 フィックスパック 1 (15.0.0.1) で実行している場合のみ表示されます。バージョン 15 では使用できません。

スコアリングされたファイルのローカル・コピーのダウンロード。 スコアリング出力をフラット・ファイルに送信するように選択した場合、必要に応じて、ファイルのローカル・コピーをダウンロードすることもできます。管理者は、ブラウザー・ベースの IBM SPSS Collaboration and Deployment Services Deployment Manager を使用して、この機能を有効または無効にできるほか、許容される最大ファイル・サイズを増減

することができます。「ファイルのダウンロード」ボタンがアクティブでない場合は、ファイル・サイズが管理者の設定した最大サイズを超えているか、または管理者がこの機能を無効にしています。この機能は、デフォルトで無効になっています。この機能は、IBM SPSS Modeler Server 15 フィックスパック 1 (15.0.0.1) でのみ使用できます。バージョン 15 では使用できません。

モデル作成用の新しいオプション設定。 IBM SPSS Modeler Advantage でモデルを作成するときに、新しいオプション設定を使用できます。これには、データ区分オプション、モデルを作成するときに含めるモデリング手法 (アルゴリズム) の選択、保持および組み合わせるモデルの最大数の指定、モデルを作成する前の利益の基準の設定などがあります。使用できるオプションは、モデルのタイプと、選択したターゲットによって異なります。詳しくは、60 ページの『オプションのモデル設定』のトピックを参照してください。以下の新機能は、IBM SPSS Modeler Server 15 フィックスパック 1 (15.0.0.1) でのみ使用できます。バージョン 15 では使用できません。

- モデル作成時に含めるモデリング手法の選択
- 利益の基準の設定
- 個々のモデルに関するハイレベル詳細情報の表示
- 選択したモデルに関する追加情報の表示

ツリー・ビューアー。 IBM SPSS Modeler Advantage でモデルを作成し、個々のモデル結果を表示した後、各モデルの名前はリンクになり、このリンクを開くと、ターゲット・フィールドや入力フィールドなど、モデルに関する情報が含まれた新規ダイアログが開きます。選択したモデルがツリー・モデルである場合、ツリー・ビューアーも表示されるようになりました。ツリー・ビューアーは、IBM SPSS Modeler Server 15 フィックスパック 1 (15.0.0.1) でのみ使用できます。バージョン 15 では使用できません。

評価の機能拡張。 IBM SPSS Modeler Advantage の「評価」ダイアログ・ボックスにさまざまな機能拡張が行われました。例えば、新しい「**モデル選択の変更**」ボタンを使用して、評価に含めるモデルを変更できます。選択を行った後、必要に応じて、現在のモデルの選択内容をコピーして個々のモデル結果テーブルに戻すこともできます。一部の新規オプションは、特定のモデル・タイプについてのみ使用できます (ほとんどの新規オプションは予測 モデルに対してのみ使用できますが、一部のオプションは予測ルール モデル・タイプにも使用できます)。「評価」ダイアログ・ボックスの以下の新機能は、IBM SPSS Modeler Server 15 フィックスパック 1 (15.0.0.1) でのみ使用できます。バージョン 15 では使用できません。

- 数値ターゲット用のパフォーマンス・グラフ
- 数値ターゲット用の利益シミュレーション
- 一致行列グラフ (すべてのターゲット用)
- すべてのグラフのテーブル・ビュー
- 「**モデル選択の変更**」ボタン
- 「**パフォーマンス・グラフについて個々のモデル評価を含める**」オプション

Microsoft Windows 8 および Internet Explorer 10。 IBM SPSS Collaboration and Deployment Services 5 フィックスパック 1 (5.0.0.1) では、Windows 8 および Internet Explorer 10 に対するサポートが追加されています。これらは、IBM SPSS Collaboration and Deployment Services 4.2.1 または 5 ではサポートされていません。

ユーザー・インターフェースの機能拡張。 ユーザー・インターフェースにさまざまなマイナー・チェンジが行われ、使いやすさが増しました。

バージョン 7

IBM Analytical Decision Management 7 (7.0.0.0) では以下の機能が追加されました。

最適化。 最適化は、マーケティング・キャンペーンからの収益の最大化や、不正や顧客離れのリスクの最小化など、具体的な目的を満たす最適なソリューションを特定することを目標にしています。最適化を行うと、最適化する価値 (収益など) を最適な方法で最大化し、母集団内で考えられるすべての組み合わせに対して、この問題に対処するためのソリューションを得ることができます。

クラスタリング・モデル。 IBM SPSS Modeler Advantage の新機能であるクラスター・モデリング手法が、類似する値またはパターンを持つデータ・レコードのグループを検出する際に役立ちます。これらの手法はマーケティング (顧客のセグメント) や他のビジネス・アプリケーションで使用されています (単一レコード・クラスターに分類されるレコードにはエラーが含まれていたり、不正の事例であったりする可能性があります)。クラスタリング・モデルはターゲット・フィールドを使用せず、true または false として評価できる特定の予測を返さないため、教師なし学習モデルと呼ばれることがよくあります。IBM Analytical Decision Management では、クラスター・モデルを自動的に作成できます。また、クラスターの結合、名前変更、変更、または削除を行ったり、以前のビジネス・ナレッジに基づいてカスタム・クラスターを追加したりすることもできます。

アソシエーション・モデル。 IBM SPSS Modeler Advantage の新機能であるアソシエーション・モデルを使用し、共に発生する可能性が最も高い項目を予測し、それらの項目間の関連の強さを予測できます。アソシエーション・ルールは、特定の結論 (特定の製品の購入など) と条件セット (複数の他の製品の購入など) を関連付けます。

データ。 「データ」タブには、複数のデータ入力値を使用する機能、「アプリケーション」ビューでプロジェクト・データ・モデルを定義する機能、データのようにモデル出力を参照する機能、ローカル・データ・ファイルをアップロードする機能、およびデータをプレビューおよびカウントする機能が組み込まれました。

IBM Cognos データ統合に対する改善。 IBM SPSS Modeler 15 と同様に、IBM Analytical Decision Management は Cognos 事前作成済みレポート、Cognos フィルターと計算値、および自動分割の制御のインポートをサポートしています。

複数の要素。 アプリケーションに、複数のディメンション階層を組み込めるようになりました。アプリケーションの起動時に、選択したアプリケーションが複数の要素を含むように構成されていると、アプリケーションで使用する 1 つ以上の使用可能な要素 (階層) を選択するように求めるプロンプトが表示されます。「定義」タブのディメンション・ツリーによって、返すことができる決定または推奨の範囲が指定されます。アプリケーションに複数のディメンション階層が含まれている場合、それらは「定義」タブのディメンション・ツリー領域の別個のタブに含まれます。

動的割り当て。 以前のリリースでは、何かを割り振る場合、それがディメンション・ツリーに存在する必要がありました。現在は、子を持たないディメンション項目の場合、プロジェクト・データ・モデル上のキーによってリンクされているテーブルから、項目を読み取ることができます。例えば、子を持たない *Cars* というディメンション項目があるとします。*BMW*、*Audi*、および *Ford* など、テーブル内のフィールドを参照するように、*Cars* に動的割り振りを設定できます。

複数割り当て。 新しい「複数割り当て」オプションを使用して、セグメントごとに複数のディメンション項目 (複数のオファーなど) にレコードを割り当てることができます。1 つのルールに 1 つの割り当てではなく、各ルールに複数の割り当てを行うことができます。

ユーザー・インターフェースの改善。 IBM Analytical Decision Management ユーザー・インターフェースに対して多数の変更と追加が行われ、すべてのアプリケーションのユーザビリティとスケーラビリティが改善されました。

ギャラリーの機能拡張。同時に複数のギャラリー項目を削除できる「ギャラリーの消去」オプションなど、ギャラリーにさまざまな機能拡張が行われました。

ルール・エディターの機能拡張。ルール・エディターを起動する場所に応じて、選択ルールのルール・タイプ (含ままたは除外) を設定し、集計ルールのリスク・ポイントを割り当て、割り当てルールのディメンションを選択し、論理和 および論理和分割 コントロールを使用して、ルール内の複数の式を組み合わせ、論理和 式を作成できるようになりました。ルールを作成するときに、階層ツリーからディメンションを選択したり (複数のツリーを持つアプリケーションの場合)、式エディターを起動して式を作成したりすることもできるようになりました。以前は、ルール・エディター・ダイアログ・ボックスで、この機能は使用できませんでした。

IBM Operational Decision Management 統合の拡張。 IBM SPSS Collaboration and Deployment Services Deployment Manager で、「新しい外部ルールの追加」ダイアログ・ボックスを使用し、IBM Operational Decision Management ルール・セットへの参照を定義して、それを IBM Analytical Decision Management で使用できるようになりました。以前は、これらのルールには Web サービスを介さなければアクセスできませんでした。このリリースでは、IBM Operational Decision Management データベースのローカルルールを参照できます。新しいローカル・オプションでは、IBM SPSS Modeler Server (IBM Analytical Decision Management によって使用されるために Deployment Manager に定義されているサーバーと同じサーバー) を介してルールが実行されます。

仮定の機能拡張。 仮定ダイアログ・ボックスの機能に、さまざまな機能拡張が行われました。例えば、シミュレーション実施日に対して、仮定評価を実行する日の範囲を入力できるようになりました。

ストリーム・ファイルのアップロード。 ローカル・ファイル・システムから任意の IBM SPSS Modeler ストリーム・ファイルを開いて、必要に応じてリポジトリに保存できるようになりました。

複数の IBM SPSS Modeler Server 資格情報。 複数の IBM SPSS Modeler Server 資格情報を IBM SPSS Collaboration and Deployment Services Deployment Manager に定義している場合、複数の資格情報を指定して IBM Analytical Decision Management で使用することもできるようになりました。そうすると、IBM Analytical Decision Management ユーザーがログオンすると、サーバーはリスト内の最初の資格情報を使用してそのユーザーが権限を持っているかを確認し、適切な権限を持つ資格情報が見つかるまで、リスト内を順番に確認していきます。順番に並べられたリスト内のいずれの資格情報に対してもユーザーが権限がない場合、そのユーザーは IBM Analytical Decision Management を使用できません。これにより、IBM SPSS Modeler Server ファイル・システムのどの部分にどのユーザーがアクセス権を持っているのかについて、より詳細に制御することができます。

バージョン 6.2

IBM SPSS Decision Management 6.2 には、以下の機能が追加されました。

Cognos データに対するサポート。 IBM SPSS Modeler Server 14.2 以降とともに使用する場合、IBM Analytical Decision Management 6.2 には、データ・ソースを定義するときに IBM Cognos Business Intelligence サーバーからデータを読み取り、スコアリング結果を Cognos BI サーバーに書き込んで戻す機能が追加されました。Cognos BI サーバーへの接続を定義した後、パッケージ、ディメンション、クエリー、およびレポートを参照して、必要なフィールドを選択できます。詳しくは、39 ページの『IBM Cognos BI ソース』のトピックを参照してください。

ILOG Business Rule Management System の統合。 ILOG などのビジネス・ルール管理システムで作成されたルールが、現在の IBM Analytical Decision Management プロジェクトで使用されているものと同じデータ・モデルをサポートするように作成されていた場合、それらのルールを IBM Analytical Decision Management アプリケーションで参照および使用できます。これらの外部ルールは、IBM Analytical

Decision Management アプリケーション内の他のモデルやルールと同じ方法で選択したり、使用したりすることができます。詳しくは、50 ページの『IBM Operational Decision Management からのルールの使用』のトピックを参照してください。現在、ILOG は IBM Operational Decision Management と呼ばれています。

Linux on System Z のサポート。 SUSE Enterprise Server 10 は、64 ビットの IBM System z10 でサポートされます。

第 2 章 顧客応答のモデリングの例

IBM SPSS Modeler Advantage

IBM SPSS Modeler Advantage は、ビジネス・ユーザーが予測モデルを作成できるようにする使いやすいアプリケーションです。予測モデルを使用して、過去に発生した事象からパターンを特定し、それらのパターンを使用して将来発生する事象を予測することができます。

例えば、モデルを使用して、収入、年齢、勤務先、会員資格などの特性を基に、解約の可能性が低い顧客や、特定のオファーに反応する可能性が高い顧客を予測できます。得られた予測は、戦略的プランニングの入力としてターゲットの顧客や対象ケースのリストを生成するために使用できるほか、予測アプリケーションとの関連でルールに統合することができます。

ダイレクト・メール キャンペーンの場合の利益の最大化

この予測モデルの例では、架空の銀行で、金融商品およびサービスの顧客へのマーケティングを担当していると仮定します。最新の試験的なキャンペーンでは、多数の顧客に複数の新規サービスの 1 つをオファーし、彼らの応答を記録しています。ただし、反応率は期待値よりも低く、キャンペーンを収益源とすることに失敗しています。

IBM SPSS Modeler Advantage の使用目的は、予想するコスト、収益、その他の要素のもとで、最適な顧客に的を絞ることで損失を利益に転換することです。

試験的なキャンペーンの結果の要約を以下に示します。

表 1. テスト・キャンペーンの結果

テスト・キャンペーンの結果	
メール数	3000
応答数	111
反応率	3.7%
コスト	6750
収益	6660
総利益（損失）	(90)

この例では、IBM Analytical Decision Management に付属してインストールされたデータ・ファイル *bank_customer_data.txt* および *bank_response_data.txt* を使用します。この例の完全なバージョンは、*bank_response_model.str* にも収録されています。インストールおよびサンプル・ファイルについての詳細は、必要に応じて、管理者に問い合わせてください。詳しくは、『サンプル・ファイル』のトピックを参照してください。

サンプル・ファイル

IBM Analytical Decision Management には、ユーザー・ガイドに示されている例の再現に使用できるデータおよびその他のサンプル・ファイルが付属しています。

データ

インストール時に、多数のサンプル・データ・ファイルが IBM SPSS Modeler Server の *Demos* ディレクトリ (例えば *C:\Program Files\IBM\SPSS\ModelerServer\16\Demos*) にインストールされます。これらのファイルはおそらく事前構築されたサンプル・アプリケーションで使用され、以下を含みます。

- *bank_customer_data.txt*
- *bank_response_data.txt*
- *customer_contact_history.txt*
- *customer_summary_data.txt*
- *insurance_claim_data.txt*
- *insurance_fraud_data.txt*
- *inventory_orders.txt*
- *inventory_stock.txt*
- *machine_log_data.txt*
- *machine_summary_data.txt*
- *retail_purchase_data.txt*

これらのデータ・ファイルを使用するサンプル・アプリケーション・ファイル、モデル、およびルールは、上記のデフォルトの場所が使用されることを前提としています。データ・ファイルが別のドライブまたは場所にインストールされる場合は、データ・ソース・パスを適宜更新する必要があります。

サンプル

既製の IBM Analytical Decision Management アプリケーションで使用できる多数のサンプル・アプリケーション・ファイル、モデル、およびルールが用意されています。サンプルは、インストール・ディスクの *Demo* ディレクトリにパッケージ化されているか、ダウンロード可能なファイルで提供されます。1 つ以上の既製アプリケーション用のサンプル・ファイルを IBM SPSS Collaboration and Deployment Services Repository にインポートするには、以下のステップを実行します。

1. **DVD** の場合は、DVD の *Demo* フォルダーにある *DecisionManagementExamples pes* ファイルを一時的な場所にコピーします。

または

ファイルをダウンロードする場合は、ファイル *SPSS_Decision_Mgmt_80_demo.zip* を一時的な場所にダウンロードして解凍します。

2. IBM SPSS Collaboration and Deployment Services Deployment Manager を開き、コンテンツ・サーバー接続にログオンします。
3. 「コンテンツ・リポジトリ」フォルダーを右クリックし、「インポート」を選択します。
4. 「インポート」ダイアログ・ボックスで「参照」をクリックし、ステップ 1 で一時的な場所にコピーした *DecisionManagementExamples pes* ファイルを選択します。次に「OK」をクリックします。
5. 「インポート競合」ダイアログで「OK」をクリックします。インポートが完了すると、IBM Analytical Decision Management で開くことのできるサンプル・プロジェクトが格納された以下の新しいフォルダーが表示されます。

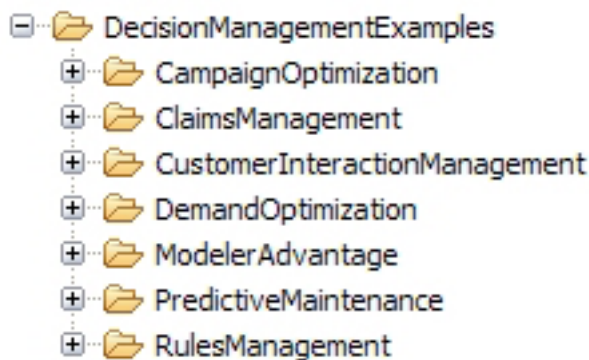


図1. サンプル成果物のフォルダー

すべてのサンプル・アプリケーション・ファイル、モデル、およびルールは、データ・ファイルが IBM SPSS Modeler Server の *Demos* ディレクトリー (例えば *C:\Program Files\IBM\SPSS\ModelerServer\16\Demos*) にインストールされていることを前提としています。データ・ファイルが別の場所にインストールされる場合は、ローカル構成用にデータ・ソース・パスを更新する必要があります。

アプリケーションの起動

IBM Analytical Decision Management 起動ページには、サイトで利用可能なアプリケーションへのリンクが表示されます。これらのリンクは管理者によって構成されますが、「アプリケーションの追加」リンクをクリックして追加のアプリケーションを選択することもできます。

1. 例の操作を開始するには、IBM Analytical Decision Managementの起動ページを開きます。後で簡単にアクセスできるようにするため、ブラウザーのこのページにブックマークを追加するか、デスクトップにショートカットを作成することができます。
2. 新規モデルを作成するには、IBM SPSS Modeler Advantage グループ内のアプリケーション・ドロップダウン・リストから「新規」を選択し、「実行」をクリックします。IBM SPSS Modeler Advantage ホーム・ページが表示されます。



図2. IBM SPSS Modeler Advantage ホーム・ページ

ホームページ上の各アイコンは、作成用のデータ定義から、モデルのスコアリングまで、モデリング・プロセス内の異なるステップを表現しています。作業を開始するには、「データ」アイコンをクリックします。

データ・ソースの設定

「データ」タブにより、例で使用される複数のサンプル・データ・ファイル用の複数のデータ・ソースを構成できます。データ・セットは、管理者が構成するか、またはユーザーが独自に追加できます。

プロジェクト・データ・モデルでは、アプリケーションで使用するフィールドを定義します。これは、ルールやモデルの評価で必要になります。他のすべてのデータセットは、アプリケーションで使用するためには、このデータ・モデルにマップする必要があります。

この例では、2 つのデータ・ファイルを使用します。これらのファイルは共にアプリケーションに付属しています。

- **bank_customer_data.txt** には、顧客 ID、返済履歴、保有する商品の数、顧客となってから経過した月数、および年齢や収入のようなデモグラフィック情報など、現在の顧客に関するデータが含まれています。プロジェクト・データ・モデルはこのファイルに基づいています。
- **bank_response_data.txt** には、以前のキャンペーンで 1 つ以上のサービスをオファーした顧客と、その応答を記録した情報が含まれています。アプリケーションは、このデータを分析することで、将来のキャンペーンで類似のオファーに対して応答する確率の高い顧客を特定します。

デフォルトでは、サンプル・データ・ファイルは、お使いのアプリケーションからアクセス可能な IBM SPSS Modeler Server Demos フォルダにインストールされます (例えば、`C:\Program Files\IBM\SPSS\ModelerServer\16\Demos`)。必要に応じて、これらのファイルのインストール支援について管理者にお問い合わせください。

プロジェクト・データ・モデルを定義するには:

1. 「プロジェクト・データ・ソース」パネルの「データ・ソース」ドロップダウンから、「データ・ソースの追加」を選択します。
2. データ・ソース名として *bank customer data* を指定します。
3. 「ファイル」を選択し、*bank_customer_data.txt*。
4. **入力フィールドの指定** の下で、不良返済レコードがあるフィールドの測定タイプが **フラグ型** になっていることを確認します。フラグの設定が、その属性の有無を示すための、真および偽を表す (一般的には *T* and *F* もしくは数値フィールドの 0 と 1 にコード化されます) 2 つの値のどちらかを持つことを意味します。
5. **入力フィールドの指定** の下で、性別フィールドの測定タイプが **名目** になっていることを確認します。(このフィールドの値は、単に、性別が男性か女性かを示すだけであり、フラグ型の場合とは異なって、**真 / 偽** を表現するものではありません。)
6. 「保存」をクリックします。

これは最初に定義したデータ・ソースであるため、プロジェクト・データ・モデル用に自動的に選択され、アプリケーションに必要なフィールドを定義するために使用されます。その他のデータ・セットはすべて、このソースに相対的にマップされます。

応答データ・ソースを定義するには:

応答データセットは、プロジェクト・データ・モデルで使用されるフィールドと同じフィールドの多くを含み、さらに 2 つのフィールド **キャンペーン ID** および **応答** を含んでいます。これらのフィールドは、各顧客に提示したオファーと、応答の有無を示します。これは、将来、応答する確率の高い顧客を予測するためのモデルを作成するときに使用されます。

1. 「プロジェクト・データ・ソース」パネルの「データ・ソース」ドロップダウンから、「データ・ソースの追加」を選択します。
2. データ・ソース名として *bank response data* を指定します。
3. 「ファイル」を選択し、*bank_response_data.txt* を選択します。

4. 入力フィールドの指定 の下で、応答 フィールドの測定タイプが**フラグ**になっていることを確認します。
5. 指定された入力フィールドの下でキャンペーン IDフィールドのとなりのチェック・ボックスの選択を解除することは、このフィールドはモデリング時には入力に使用しないことを示します。（このフィールドは、ある顧客が過去にどのキャンペーンに選ばれたかを示すために使用することができますが、将来の応答を予測するには不便です。）
6. 「保存」をクリックします。

データモデル内のいくつかのフィールドは応答データセット内には含まれないため、「データ」タブの **互換** 列の下でのこのデータ・ソースにはオレンジ色のアイコンが付けられており、マッチするフィールドはあるが、すべてのフィールドがマッチしているわけではないことが示されます。このデータ・セットはモデリングには適していますが、シミュレーション、スコアリング、または導入には適していません。（これらの目的で使用するには、データ・モデルのすべてのフィールドが揃っている必要があります。）

モデルの構築

「モデリング」タブでは、データ・ソース、ターゲット・フィールド、およびその他のオプションを指定できます。

1. データソースとして、*bank response data*を選択し、ターゲットフィールドとして、「応答」を選択します。
2. **データスキャンが必要です**というメッセージが表示されたら、アイコンをクリックしてデータスキャンを行います。次に「**モデルの作成**」をクリックします。

モデルが完成すると、分布図が表示されます。分布図では、観測値が横軸に表示され、値 1 が見込みのある応答を示します。

- 棒の高さは、反応率を反映します。応答データセット内の 3,000 人の顧客のうち、111 人が応答しているため、ヒット率は 3.8% です。
- 各棒内の色付けされた領域は、モデルが予測した応答を表現しています。すなわち、右上の凡例に示したように、赤は見込みのある予測を示し、青は見込みのない応答を示します。反応率が低いため、予測した応答（赤の領域）の数は多くありません。しかしながら、「偽」のバー（応答=0）は圧倒的に青になっています。これは、応答しないという予測の圧倒的多数が、実際に応答しなかったことを示しています。

如何にモデルが実行されるかについての良いアイデアを得るには、次のセクション（『モデルの評価』）に進みます。

モデルの評価

「モデリング」タブで、「評価」をクリックし、モデルを顧客データに適用したときの動作を調べれば、モデルが最終損益に与える影響を理解することができます。

1. テスト・データ・ソースに、*bank response data*を選択します。
2. 作成するグラフ タイプに、「**差益グラフ**」と「**利益グラフ**」を選択します。
3. 「**最大利益のシミュレート**」を選択し、母集団として3000を指定します。これはテスト・キャンペーンの数と一致します。
4. コストに2,25を、収益に60を指定します。これらの値は、各オファーのコスト（制作、処理、郵送費など）と、肯定応答による期待収益を示します。
5. 「**求める反応**」で 1 を選択します。

6. 「実行」をクリックします。

差益グラフには、モデルを使用することで取得した改善または「上昇分」が示されます。モデルの作成時に（「オプション」設定の下で）「データを自動的にパーティション分割」オプションを選択していたため、トレーニング・パーティションとテスト・パーティションには異なるグラフが表示されます。各グラフには、サンプル内のヒット総数と比較した、各パーセンタイルにおけるヒット（見込みのある応答）数が示されます。斜めの線は、モデルを使用しなかった場合の期待される回答を作用したものです。半分のヒットで良ければ、半数の顧客に連絡を取れば良いことを示しています。曲線は、モデルを使用した場合の改善を示しています。反応率の高い60%の顧客に絞ることで、メールの数を減らすにも関わらず、ほぼすべてのヒットを獲得できることが示されています。

差益グラフは、指定したコストと収益の値に基づいて、総利益の観点から差益を示します。モデル学習に使用した例では、30番目の百分位のすぐ下で収益が最高になります。

この機能を理解するには、「シミュレート」タブをクリックして、指定した母集団のサイズに対する期待最大利益を調べます。期待利益は、次の式を使用して算出されます。

$$\text{期待利益} = (\text{反応率} * \text{収益}) - \text{コスト}$$

ここで反応率は、モデルに基づいて各顧客ごとに決定されます。

モデルによって反応率が最も高いと判定された顧客にターゲットを絞ることにより、顧客あたりの収益を最大化することができます。含まれる顧客数を増やせば、ある一定の点までは総収益も増加しますが、反応率の低い顧客が多く含まれるようになると、顧客ごとの収益は徐々に減ります。シミュレーションは、最大利益を獲得するために線引きをする場所の決定を支援します。

「利益の比較」には、この結果の詳しい情報が表示されます。

- **対象を 48% に絞る。** この列には、モデルによって予測された反応率の高い 48% の顧客を対象を絞った場合の結果が表示されます。反応率が最も高い 1290 人の顧客を選択的に含めることにより、収益を最大化することができ、その結果、期待される収益は \$3277.50 になります。
- **無作為に 48%。** 同じサイズの無作為のサンプルを抽出した場合の結果を表示します。この場合、送付するオファーの数（1290）は同じですが、モデルによって提供される付加情報を使用していないため、期待される回答数は少なくなります（テスト・キャンペーンの率に一致）。その結果、全体の反応率はオファー送付コストを回収できるだけの十分な値とならず、損失を発生させることとなります。
- **すべての顧客。** 3000 名の顧客をすべて含めた場合の結果を示します。この場合、テスト・キャンペーンの結果と正確に一致します。ここでも、モデルの有益な情報を利用していないため、損失が発生しています。

結果を見ると、モデルを使用した場合に反応率が改善され、損失を生みだしていたキャンペーンを利益を生むキャンペーンに転換できることがわかります。\$90.00 の損失と \$3277.50 の利益との差が、予測モデルによって得られる利点を表しています。この収益は、テスト キャンペーンで使用した 3000 人の顧客をベースにしたものですが、対象の人数をもっと増やすことで、それだけ受ける恩恵も大きくなります。

7. シミュレーションの結果をレビューしたら、「予測結果」グラフの下の「スコアリングに使用」を選択し、指定されたパーセンテージ（この場合は 48 %）に基づいて「スコア」タブの設定を更新します。詳しくは、16 ページの『モデルのスコアリング』のトピックを参照してください。
8. 必要であれば、プロジェクトを `bank_response_model.str` として保存します。次のセクションで、モデル・タイプを変更し、新規プロジェクトを作成します。

予測ルール・モデル

予測されたモデルは、必要とするすべての予測機能を提供しますが、場合によってはモデルの内部的な動作に関してもう少し自由にコントロールしたり、仕組みを理解したいと考えることがあるかもしれません。そのような場合には、予測ルール（対話型）モデル作成を使用できます。予測ルールモデルで、データをセグメントする為にマニュアルでルールを作成出来たり、或いは選んだフィールドの本当の値の高低いずれかの集中があるセグメントを自動的に見つけられるようになります。例えば、キャンペーンに対して肯定的な反応を示した顧客を探して、より反応確率が高いセグメントを割り出します。

1. モデルタブに戻って、**モデルを変更**をクリックして**予測ルール・モデル**を選択します。
2. データソースとして、*bank response data*を選択し、ターゲットフィールドとして、「**応答**」を選択します。
3. 「求める反応」に、**1**を指定します。これは、見込みのある応答を行った顧客を検索することを意味します。
4. 「**作成の開始**」を選択します。
5. 「**モデルの拡張**」を選択し、より高い反応率を持つセグメントを特定します。最高の反応率のセグメントを特定するための3つの最善のルールが、作成され、表示されます。
6. **包含/除外** アイコンをクリックして、各ルールによって選択された顧客の包含または除外を選択します。たとえば、2番目と3番目のルール・セグメントは反応率が8%強です。これはキャンペーン全体の3.7%よりは高率ですが、1番目のセグメントの15.6%ほどは高くありません。これらの2つのルールの包含または除外が、全体の収益にどのような影響を与えるを見るためには、両方の収益シミュレーションを行うことができ、その結果を以下の通り比較します。
7. 3つのセグメントをすべて**包含**に設定し、「**評価**」をクリックします。
8. 作成するグラフタイプに、「**差益グラフ**」と「**利益グラフ**」を選択します。
9. 「**最大利益をシミュレート**」を選択して母集団に「**3000**」を、コストに「**2.25**」、そして収益に「**60**」を指定します。これらは、予測モデルで評価した際に使用した仕様と同じです。「求める反応」には自動的に値が入力されます。
10. 「**実行**」をクリックすると、グラフの作成が完了した時点で、「シミュレート」タブへ移動します。

この場合、このモデルでシミュレートされた最大収益は \$3178.90 で、自動化されたモデルの場合よりもわずかに低くなります。必要に応じて、「評価」ダイアログを閉じて、2番目および3番目のセグメントの設定を**除外**に変更し、収益シミュレーションを再実行して結果がどのように変わるかを確認してみてください。予測ルール（対話型）モデル作成の利点は、ユーザーが動作をコントロールし、理解を深めることができる点にあります。

11. 必要であれば、プロジェクトを *bank_predictive_rule_model.str* として保存します。

モデルのテスト

モデルの作成が完了したら、一部の顧客レコードをテストして、キャンペーンへの反応率を調べることができます。

1. モデルのスコアリングによる特定のレコードの処理をプレビューするには、「モデリング」タブで「**テスト**」を選択します。「テスト」ダイアログが表示されます。
2. 「**レコード**」領域で、「**利益**」の数値を 100 に増やします。
3. 「**表示フィールドを選択**」領域で、「**顧客 ID**」を選択します。また、*\$XF-Response* フィールドと *\$XFRP-Response* フィールドも含まれます。これらのフィールドは、各顧客の予測される応答（この場合、0 は「いいえ」、1 は「はい」）と傾向スコアを示します。

必要に応じて、「年齢」、「最後のオファーから経過した週の数」、「不良な支払履歴あり」などのフィールドを追加することもできます。これらのフィールドは、特定の顧客がモデルにより一定のスコアリングをされる理由と仕組みを理解するのに役立ちます。

モデルのスコアリング

モデルのスコアリングとは、モデルをデータ・サンプルまたは関心のある母集団に適用することです。たとえば、最も反応する可能性のある顧客を予測するモデルを生成して、キャンペーン参加を見込める将来的な顧客をスコアリングすることができます。「評価」ダイアログボックスで、利益を最大化するためにスコアリングを行うレコードのパーセンテージが決定されました。そして、この結果に基づいて、スコアリングするレコードのパーセンテージを指定しました。

1. 「スコア」タブに移動します。
2. 「スコアリングするデータ・ソース」で、*bank response data* が選択されていることを確認します。
3. スコアリング結果の送信先を指定します (単にスコアをプレビューしたいだけの場合は、このステップはスキップしてください)。
4. 「出力フィールド」で、「顧客 ID」フィールドが選択されていることを確認します。
5. 「スコア」設定は、モデルの評価時に実行した利益シミュレーションと同じレコードのパーセンテージをスコアリングするように既に構成されています (シミュレーションの結果を参照する際に「スコアリングに使用」ボタンを選択しなかった場合は、設定をここで更新することができます)。
6. ダイアログの最下部にある「プレビュー」をクリックして、スコアリングで得られた結果のサンプルを表示します。
7. 「データ概要」をクリックして、スコアリングで使用されるデータの概要を確認します。詳しくは、32ページの『データ概要』のトピックを参照してください。

傾向に基づかないスコアを選択したので、最も応答の傾向が大きい顧客のみが含まれています。顧客 ID を顧客データベースにリンクし直すことにより、この結果を基にしてメーリング・リストを生成することができます。

8. 「スコア」をクリックします。モデルが顧客データに適用され、ダイアログの下部に確認メッセージが表示されます。

クラスター・モデリングの例

クラスターリングモデルで、類似の特徴のグループごとに記録を分類できます。これは、データ内の自然発生的なグループを識別するために役立ちます。例えば、人口統計による特徴や購買行動に基づいて顧客をセグメント化する場合です。

この例は *bank_customer_data.txt* という商品に付属してインストールされていたデータファイルを使用しています。この例の完全なバージョンは、*bank_cluster_model.str* にも収録されています。インストールおよびサンプル・ファイルについての詳細は、必要に応じて、管理者に問い合わせてください。詳しくは、9ページの『サンプル・ファイル』のトピックを参照してください。

1. アプリケーション起動ページに戻って新しい IBM SPSS Modeler Advantage プロジェクトを作成します。
2. データタブにて、プロジェクトデータソースパネルで、11ページの『データ・ソースの設定』で解説されている手順で *bank customer data* という新しいデータソースを追加します。
3. モデリングタブにて、**モデルを変更**をクリックして、**クラスターリングモデル**を選択して、**保存**をクリックします。

4. テスト・データ・ソースに、*bank customer data*を選択します。
5. 「オプション設定」セクションを展開します。クラスタリングモデルでは、評価フィールドの選択、或いはクラスタリングの必要な範囲の設定の為に、追加の**自動クラスタオプション**のオプション設定が使用可能になります。

クラスタリング・モデルが作成された時、場面の後ろでは、いくつかの異なるクラスタリングモデルが実際には作成されます。これらのモデルはその後品質の尺度によって比較、そしてランクされ、最高モデルは自動クラスタ・モデルの使用の為に選択されます。よってこれらの2つの自動クラスタ・オプション設定はどのモデルが実際に選択されているのか制御できるようにしてくれます。評価フィールドがここで指定される場合は、評価フィールドの値を上手に識別するモデルが選択されます。そして必要な範囲が設定されたら、指定された範囲外にクラスタを発見するモデルは排除されます。

6. この例として、**発見されたクラスタの必要な範囲を設定**オプションを選択して最低で**3**、最高で**15**のデフォルトを使用します。これでクラスタリング・モデルが持つクラスタの数が少な過ぎたり、多過ぎたりする事がないようにします。
7. 残りのオプション設定は他のモデリング・タイプのもと同じです。**信頼できるモデル作成のために自動的にデータをクリーンアップ**オプションの選択を取り消して、ただ自動クラスタリングの例を簡単に理解できるようにします。
8. **オプション設定**パネルを省略して**クラスタ・パネル**を展開します。

クラスタ・パネルには、手動クラスタのセクションと自動クラスタのセクションがあります。手動クラスタはデータの知識に基づくクラスタの定義を可能にしてくれます。この例として、手動クラスタを高価値顧客に、そして手動クラスタを若い独身の顧客に、次のようにして定義します。

9. 「**新規のルールを作成**」アイコンをクリックします。ルールを高価値顧客という名前にして、次を式に追加して、「**OK**」をクリックします。

顧客としての月>12

収入>65000

10. 次の式で若い独身の顧客という新しい別のルールを作成します。

年齢<35

結婚暦=独身

11. **手動クラスタ**セクションの下に2つの新規ルールがリストされます。これらの手動クラスタに何人の顧客が捕らえられているかを知るには、**クラスタレコードカウント**をクリックします。その後手動クラスタを再度展開して、カウント列を見て各クラスタに何人の顧客が見つかったかを確認します。67 人の高価値顧客と 33 人の若い独身の顧客が確認されたことがわかります。
12. 次に**自動クラスタ検出**をクリックして関心のある他のクラスタを自動的に検出します。終了したら、「**自動クラスタ**」セクションを展開します。自動クラスタが 3 つ確認されたことがわかります。
13. **自動クラスタを見る**をクリックして更に詳細を見ます。自動クラスタ結果ビューアーが表示されます。

モデル要約ページは使用された自動クラスタモデルのアルゴリズム（この例では*TwoStep*）に関する基本情報、検出された自動クラスタの数、そしてクラスタの品質の基本的な考えを提供します。

他のグラフも使用可能です。各グラフの様々な区域をうろついて更なる詳細を確認したり、いくつかのグラフ（例えば予測の重要度グラフ）と相互に影響できます。

クラスタービュー（一番最近のグラフ）を使用して、クラスターを比較して興味深いパターンを探し出す事も出来ます。現実世界の実行においては、自動クラスターを研究してその後より意味のある名称（クラスタービューの各クラスターラベルをクリックして）与えます。クラスター・ラベルの名称を更新したら、新しい名称がすべての他のグラフにて使用され、モデリング・タブに戻ります。次に 評価に進んでモデルをテストし、結果に満足できれば最後に採点します。

14. 必要であれば、プロジェクトを *bank_cluster_model.str* として保存します。

クラスター・モデルの詳細については、57 ページの『クラスタリング・モデルの作成』を参照してください。

アソシエーション・モデリングの例

アソシエーション・モデリングは予測分析における重要な技術です。アソシエーション・モデルを使用すると、一緒に出現する可能性が最も高い項目を予測し、それらの項目間の関係の強度を予測できます。例えば、顧客が現在ショッピング・カートに入れている商品に基づいて、次回に購入する可能性が最も高い商品を割り出します。

この例は *bank_purchase_data.txt* という商品に付属してインストールされていたデータファイルを使用しています。テレビ、パソコン、そしてスマートフォンといった電化製品に関する購入情報で構成されているデータ。

この例の完全なバージョンは、*retail_association_model.str* にも収録されています。サンプル・ファイルのインストールの詳細については、必要に応じて、管理者に問い合わせてください。詳しくは、9 ページの『サンプル・ファイル』のトピックを参照してください。

1. アプリケーション起動ページに戻って新しい IBM SPSS Modeler Advantage プロジェクトを作成します。
2. データタブにて、プロジェクトデータソースパネルにて、*retail_purchase_data* という新しいデータソースを追加します。「ファイル」を選択し、*retail_purchase_data.txt* を選択します。
3. モデリングタブにて、**モデルを変更**をクリックして、**アソシエーションモデル**を選択して、**保存**をクリックします。
4. テスト・データ・ソースに、*retail_purchase_data* を選択します。
5. テスト・データ形式に、「**テーブル形式**」を選択します。トランザクション・データの設定は含まれたアイテムを表示するだけでなく、すべてのアイテムが単一の列に現れます。しかしこのデータセットにおいては、各アイテムは、購入されたかそうでないかを表示する真、或いは偽のフラグのある別の列にあります。
6. 「**モデルの作成**」をクリックします。製作に失敗しました、しきい値が高すぎますというエラーメッセージを受け取ります。エラーを閉じる。

アソシエーション・ルールを求める時は、一般的に適用可能、信頼できる、そして簡単なルールを捜し求めることを通常はお勧めします。

7. 「オプション設定」セクションを展開します。「アソシエーション・ルール・オプション」の下で、最初のしきい値を 10 から 1 に変更し、2 番目のしきい値を 80 から 1 に下げます。通常、異なるしきい値で実験し、そのたびにモデルを再構築して、ルールが多過ぎたり、足りなかったりしないようにバランスを取ります。
8. 「**モデルの作成**」をクリックします。

モデルの結果において、各行は別個のアソシエーション・ルールを表現しています。リストの一番最初のルールは顧客がスマートフォンを購入する場合、大画面テレビも合わせて購入する場合がありますと説明

します。「**範囲 (%)**」列は、15% (つまり、顧客 6 人につき約 1 人) がスマートフォンを購入したことを示しており、「**確信度 (%)**」列は、それらの顧客のうち、大画面テレビも購入する確率が約 30% であるということを示しています。

よって、少数のルール (この例では239) でも、ビジネス・コンテキストから次に顧客が何を購入するかを予測するために役立つ可能性があります。

- 次に、**テスト**をクリックしてその後**実行**をクリックして予測や推薦がどのように成されるのかを確認します。

この場合、モデルは最高確信数値のルールを見て、ユーザが既に購入したのを見て、そしてその後これら2つを元に顧客が何を購入するかの推薦を行います。

- 最初の記録の横側にあるプレビューアイコンをクリックして、例えば大画面テレビとスピーカーを最初に購入したけど、他の商品はまだ何も購入していない顧客を確認したりします。プレビューを閉じる。

この顧客への推薦は2列目に確認できるように標準テレビの購入である。3 列目の値 **0.3** は、大画面テレビとスピーカーを既に購入している顧客が標準テレビも購入する可能性は 30% であるという意味です。

しかしながら、これが例外なのか一般的に適用可能な予測されたトレンドなのかをどのように判別するのか?ルール番号 (この例では「**115**」) を記録することで、戻ってルールの更なる詳細を確認できる。

- テスト・ダイアログを閉じる。
- 115という数字を確認するまでモデルの結果をページをめくります。なお、多数の結果ページがある大きなモデルでは、「**ID による規則の検索**」をクリックしてルールの番号を入力することもできます。
- 結果列のヘッダーの横にあるいずれかの矢印をクリックして、以下の項目を選択します。

「列」 > 「**範囲 (N)**」

「**範囲 (N)**」列が表示されます。事例数によると、ルール 115 が 3 つの商品を一緒に購入した 58 人の顧客に基づいているので、これが例外ではないということが明らかになりました。これは我々が将来の推薦の基本として使うことに十分な自信を与えてくれます。

どうやって推薦に使用できるかを確認する単一のルールを見ています。しかしIBM Analytical Decision Managementはまた、モデルを我々の特定の目的に、より関連付けるようにカスタマイズする強力なオプションも提供します。例えば、どの顧客が大画面テレビの購入に関心があるのかについて予測したいかも知れない。

- 「**適用**」をクリックします。これはフィルターを定義して、フィルターに一致しているルールのみがモデルに表示されるダイアログ・ボックスを表示します。
- 「**フィルターの有効化**」を選択して、アイテムに大画面テレビを選択します。
- 「**保存**」をクリックしてモデルの結果に戻ります。

ルールの番号は 239 から 34 に減少しました。

適切な数の顧客に基づいたルールのみを使用していることを確認したい場合もあります。これを達成するにあたって、10 人以下の修正に基づいたルールを除外するかもしれません。

- 「**範囲 (N)**」列のヘッダーの隣の矢印をクリックし、「**昇順で並び替える**」を選択します。
- 「**除外**」列で、「**範囲 (N)**」の値が 10 未満のすべてのルールを選択します。

19. この時点で多くのルールがフィルターされて除外されたとはいえ、まだモデル内にあるということに注意して下さい。モデルのスコアリング時に使用されないように永久に削除するには、「除外およびフィルタリングされたルールの削除」をクリックし、「OK」をクリックします。

返還されたルールのは数は10、或いはそれ以上の事例全部、そして大画面テレビを予測した全てを元にした20にまで減少した。

20. 適用可能なルールを確認するには、**テスト**をクリックして、その後**実行**をクリックします。大画面テレビは今ではさらに数名の顧客に推薦されていて、また何人かは他よりも大きな自信を持っています。
21. テスト・ダイアログを閉じる。
22. オプション設定セクションでは、スコアリング・オプションのセクションを展開してアソシエーション・モデルに特定したスコアリング・オプションを表示します。

デフォルトでは、予測の最大数は 3 です。これは最大確信数値に基づきます。あるいは、範囲、ルール・サポート、最大数のリフト (購入される特定のアイテムの確率の上昇)、または展開性に基づいてルールを選択するようにルール基準を変更できます。また「予測を表示させない」を選択することにより既に1台購入済みの顧客には、大画面テレビをオファーしないようにしたい知れません。

デフォルト・スコアリング・オプションはほとんどの状況に適しています。しかしいくつかの場合において、これらのオプションが提供する追加の柔軟性は有利に働くかも知れません。

このプロセスにおける次の段階はスコアリング・タブを使ってモデルを判定します。詳しくは、69 ページの『データベース・テーブル、ファイル、Analytic Server、または、IBM Cognos BI サーバーのモデルのスコアリング』のトピックを参照してください。

23. 必要であれば、プロジェクトを *retail_association_model.str* として保存します。

アソシエーション・モデルの詳細については、58 ページの『アソシエーション・モデルの作成』を参照してください。

第 3 章 作業領域のナビゲート

表示言語の変更

各ユーザーは自分のアプリケーションの言語をサーバーも再起動させることなく簡単に変更することができます。例えば、ドイツを話すユーザー、英語を話すユーザー、そして、日本語を話すユーザーは自身の言語で同じアプリケーションを使用することができます。

使用するブラウザの「ツール」メニューから、「インターネット・オプション」または「オプション」を開いて、使用したい言語に切り替えて画面をリフレッシュします。使用したい言語に IBM Analytical Decision Management が提供されない場合、デフォルトでは英語、または存在する場合は第 2 言語で表示されます。

IBM Analytical Decision Management インターフェースでは以下の言語がサポートされます。このリストはリリースごとに変更される可能性があり、製品資料の記載は各言語のユーザー・インターフェースと必ずしも一致しない場合があります。

- 英語
- ドイツ語
- スペイン語
- フランス語
- イタリア語
- 日本語
- 韓国語
- ポルトガル語 (ブラジル語)
- 中国語 (中国/簡体字)
- 中国語 (台湾/繁体字)

アプリケーションの起動

「アプリケーション起動ページ」は、インストールされているアプリケーションへの構成可能リンクを提供します。各パネルはインストールされているアプリケーションを表します。ドロップダウンをクリックして、最近保存されたプロジェクト、モデル、またはルールから選択します。ラベル付けされた特定のバージョンを指す 1 つ以上のカスタム・ショートカットを追加することもできます。

- 「アプリケーション起動ページ」にアクセスするには、<http://hostname:port/DM> にナビゲートします。ここで、*hostname* は、リポジトリ・マシンの IP アドレスまたは名前であり、*port* は、アプリケーション・サーバーのポート番号です。この情報が不明である場合は、管理者に問い合わせてください。また、ブラウザのこの画面をブックマークに追加したり、デスクトップにショートカットを作成することもできます。

IBM Analytical Decision Management URL または IBM SPSS Collaboration and Deployment Services URL のいずれかにアクセスする際に問題が発生する場合は、管理者がそれらをカスタマイズしている可能性があります。管理者にお問い合わせください。詳しい情報については、*IBM SPSS Collaboration and Deployment Services* リポジトリのインストールおよび設定ガイドの「アプリケーションのコンテキスト・ルート」の章を参照してください。

- 新規のプロジェクト、モデル、またはルールを作成するには、目的のアプリケーションのドロップダウン・リストから「新規」を選択し、「実行」をクリックします。
- 保存されているプロジェクト、モデル、またはルールを開くには、アプリケーション・パネルのドロップダウン・リストから最近使用されたバージョンまたはショートカットを選択するか、ドロップダウン・リストから「参照」を選択して別のオブジェクトまたはバージョンを選択します。リスト内の項目のリポジトリの場所を迅速に確認するには、ドロップダウン・メニューでその項目の上にマウス・ポインターを移動します。
- ドロップダウン・リストで常に使用可能となる新規カスタム・ショートカットを追加するには、アプリケーション・パネルの右上隅にある三角形の矢印アイコンをクリックします。オプションとして、表示するショートカットの数を指定できます。

アプリケーションを開くための別の方法

- ブラウザーのアドレス・バーに URL をコピーまたは入力するか、ブラウザの保存されているブックマークまたはお気に入りのリストに URL を追加します。
- URL を指すデスクトップ・ショートカットを作成します。
- 会社のイントラネット・サイトまたはポータル内からアプリケーションの URL にリンクします。

起動ページのカスタマイズ

「アプリケーション起動ページ」の右上隅にあるコントロールを使用すると、アプリケーションの追加と、パネル・ビューとリスト・ビューの表示の切り替えを行うことができます。また、パネルをドラッグして、ページ内で並べ替えることもできます。

- 新規アプリケーション・パネルを追加するには、「アプリケーション起動ページ」の右上隅にある「アプリケーションの追加」を選択します。(使用可能な追加アプリケーションがない場合は、このリンクは使用できません。)

使用可能なアプリケーションのリストが表示されます。オプションとして、1 つ以上のラベル付けされたバージョンへのショートカットを作成できます。

- 起動ページ上のアプリケーション・パネルの配置を変更するには、任意のパネルのタイトル・バーをクリックし、マウス・ボタンを押しながらパネルを新しい場所にドラッグします。
- アプリケーションを削除するには、アプリケーション・パネルの右上隅にある削除アイコン (赤色の X) をクリックします。そのアプリケーションを指すショートカットもすべて削除されます。(削除アイコンは、管理者によってアプリケーション・パネルがページ上でロックされている場合は使用できません。)
- 保存されていないプロジェクト、モデル、およびルールが、起動ページのギャラリーにリストされることがあります。このギャラリーには、まだ作成中のモデルや、作成は完了したがまだ保存されていないモデルが表示されます。詳しくは、23 ページの『ギャラリー』のトピックを参照してください。

ホーム・ページ

アプリケーションには、段階的なワークフローが用意されており、これはホーム・ページ上のアイコンで表されます。任意のアイコンをクリックすると、そのステップにジャンプします。

データ。アプリケーションが使用するデータ・セットを定義します。

モデル。分析データを使用してモデルを作成し、予測の重要度、分布、ゲイン・グラフなどを使用してそれらのモデルを評価することで、データに潜むパターンを発見して洞察を獲得し、それらの洞察を使用して決定の精度を上げます。

スコア。モデルをスコアリングします。

ギャラリー

ギャラリーは、「アプリケーション起動ページ」に表示されます。夜間に作成するためにそのままにされているモデルなどの、保存されていない作業（退社前にモデルを開始し、ブラウザを閉じ、翌朝戻って結果を確認することができる作業）がリストされます。また、ブラウザを誤って閉じた後に、ギャラリーを使用して、作業を取得することもできます。ギャラリー内の項目には、アプリケーションに対応するプロジェクト、ルール、モデル、データ・スキャン、評価、および他の長時間実行中のタスクが含まれることがあります。

- 項目の名前に下線が付けられている場合は、名前をクリックすると、関連項目が開きます。例えば、保存されていないデータ・スキャンがある場合、ギャラリー内のデータ・スキャン・タイトルをクリックすると、「データ」タブが開きます。
- ギャラリー内に表示されている項目で不要なものがある場合は、削除アイコンをクリックしてその項目を削除できます。その項目を完全に削除することを確認する警告メッセージが表示されます。あるいは、複数のギャラリー項目を同時に削除するには、「**ギャラリーの消去**」をクリックします。「ギャラリーの消去」ダイアログ・ボックスが表示され、すべてのオブジェクトを削除するタブを選択できます。
- ギャラリーの「プロジェクト」タブで、「**作成場所**」ドロップダウン・メニューを使用することで、各アプリケーションのタブに表示される項目をフィルターに掛けることができます。例えば、あるアプリケーションで作業中に、他のアプリケーションのギャラリー項目を表示したくない場合があります。

設定の指定

IBM Analytical Decision Management を必要に応じてカスタマイズするために調整できる設定は数多くあります。このカスタマイズは、作業の作成、テスト、および保存時に、同じアクションを何度も選択する手間を省くために、主に独自のデフォルト・オプションを設定することによって行います。指定する設定は、そのユーザーに固有のものであり、そのユーザーが使用するすべてのアプリケーションに適用されます。

ユーザー設定にアクセスするには、任意のアプリケーションから、ツールバーの「**ユーザー設定の変更**」アイコンをクリックします。



図 3. 「ユーザー設定の変更」アイコン

以下のセクションで示す説明に従って、各領域について独自の優先設定を指定できます。

一般設定

アプリケーション・ショートカット内の最近開かれたオブジェクトのデフォルト数。 アプリケーション起動ページのドロップダウン・リストから選択できる項目の数を選択します。

日付の形式。 日付を表示または記録するときに使用する形式を選択します。

時刻の形式。 時刻を表示または記録するときに使用する形式を選択します。

通貨。 各新規プロジェクトのデフォルトの通貨を選択します。別の通貨として指定された既存のデータが、この設定によって上書きされることはありません。

小数点以下の桁数。 数値に表示する小数点以下の桁数を選択します。

プレビューに表示する行の数。 プレビューを表示するとき (インポート時にデータ・ソースを確認する場合など) に表示するレコードの最大数を選択します。

ページごとの結果の数 (ページ分割するテーブルの場合)。 テーブルの表示時に各ページに表示する結果の最大数を選択します。

グラフに表示する項目の最大数。 表示する項目が多数ある場合は、グラフの結果が正しく機能しないことがあります。グラフに表示する項目の最大数を指定します。これにより、項目数がこの最大値を超えた場合に、グラフの代わりにテーブルが表示されるようになります。この設定は、データ概要のカテゴリ・フィールドと、カテゴリ・ターゲットが指定されたモデルのスコアリング要約レポート・グラフにのみ適用されます。

グラフおよびテーブルに適用するデータ・セグメント (分位) の数。 アプリケーション全体でグラフおよびテーブルに使用するデータ・セグメントの最大数を選択します (該当する場合)。

モデルの設定

デフォルトで作成するモデル・タイプ。 使用可能なデータからデフォルトでアプリケーションによって自動的に最良のモデルを作成するかどうか (予測モデル)、特定の結果と比較して可能性が高いまたは低いセグメントを識別するモデルを使用するか (予測ルール・モデル)、レコードを特定数のクラスターにカテゴリ分類するモデルを使用するか (クラスタリング・モデル)、または共に出現する可能性が最も高い項目を予測し、それらの項目間の関係の強度を予測するモデルを使用するか (アソシエーション・モデル) を指定します。このドロップダウン・メニューで選択できるのは、アプリケーションで表示するように構成されているモデル・タイプのみです。

自動的にデータを区分して、検証およびテストのためのデータ・ソース作成時にモデル評価を実行できるようにする。 このオプションを選択すると、モデルのトレーニング用とテスト用に、データが別々のサブセットまたはサンプルに分割されます。モデルの作成に使用したサブセットとは別のサブセットでテストを行うことで、その動作を他のデータ・セットに一般化する方法を推測できます。無作為に生成されるトレーニング・データ区分のサイズ とテスト・データ区分のサイズ にパーセント値を指定することもできます。これらのオプションは、アソシエーション・モデルを除くすべてのモデル・タイプに使用できます。

自動的にデータをクリーンアップして準備し、モデル作成の信頼性を確保する。 このオプションを選択すると、多くのデータ準備作業をアプリケーションで実行することができます。例えば、データに関する問題の修正、利用価値の低いデータの選別、新規属性の派生などです。

クリーンアップ操作で、処理速度と精度を同等に重視するか、処理速度よりも精度を優先するか、または非常に大きなデータ・セットを処理する場合や結果が早く必要な場合などに、処理速度を優先するかどうかを選択できます。

単一のモデルの作成に費やす時間の制限。 このオプションを使用すると、各モデルの作成に費やす時間のデフォルトの長さを指定できます。これは、大きなデータ・セットを処理する場合に特に有効で、モデル作成の完了に必要な時間に大きく影響する場合があります。

テストおよびシミュレーションの設定

テストするレコードのデフォルト数。 テストの実行時に対象となるレコードのデフォルト数を指定します。

プロジェクト、モデル、およびルールのオープンと保存

IBM SPSS Collaboration and Deployment Services Repository に対してオープン、削除、または保存操作を実行したり、ローカル・ファイル・システムに対してダウンロード操作を実行したりすることができます。IBM SPSS Modeler のストリーム・ファイルをローカル・ファイル・システムから IBM SPSS Collaboration and Deployment Services Repository にアップロードして保存することもできます。IBM SPSS Collaboration and Deployment Services Repository に保存することで、複数バージョン、ラベル付け、およびセキュリティに対するサポートなど、重要な利点を得ることができます。ローカル構成によっては、一部のオプションがサポートされない場合があります。



新規モデルを作成します。



現在のプロジェクトを IBM SPSS Collaboration and Deployment Services Repository 内の最新の既知のロケーションに保存します。オプションで、保存時に新しいラベルを適用できます。



現在のプロジェクトを保存してラベルを付けるか、新しい名前で作成します。該当する場合は、IBM SPSS Collaboration and Deployment Services Repository に保存するか、ローカル・ファイル・システムにダウンロードするかを選択できます。



IBM SPSS Collaboration and Deployment Services Repository またはローカル・ファイル・システムのうち該当する方からプロジェクトを開きます。または、該当する場合に、ローカル・ファイル・システムからプロジェクトをアップロードすることもできます。ファイル・システムへのアクセスは、すべてのアプリケーションでサポートされているわけではありません。IBM SPSS Modeler ストリームをアップロードする場合は、ファイルを開くか、リポジトリにファイルを保存するかを選択できます。

IBM SPSS Collaboration and Deployment Services Repository への保存

IBM SPSS Collaboration and Deployment Services Repository には、複数バージョン、ラベル付け、およびセキュリティに対するサポートなどの重要な利点があります。ユーザーが保存を実行するたびに、新しいバージョンが作成されます。以前のバージョンへはいつでも戻ることができます。また、ラベルを使用してバージョンを管理し、セキュリティ設定を適用して、オブジェクトのアクセス、表示、または削除を実行できるユーザーを決定することもできます。

左側ペインのツリー表示には、表示権限のある IBM SPSS Collaboration and Deployment Services Repository 内のフォルダーが表示されます。このフォルダー・ツリーをナビゲートするか、検索を実行して、目的のオブジェクトを探します。左下ペインに表示されるファイル・リストを、ラベルを基準にフィルタリングすることもできます。例えば、「実動」ラベルをクリックして、そのラベルの付いたファイルだけを表示することができます。

ファイルの詳細を表示するには、ファイルの横にあるプレビュー・アイコンまたは情報アイコン (使用可能な場合) をクリックします。

ファイルを削除するには、そのファイルの横の削除アイコンをクリックします。この操作は、IBM SPSS Collaboration and Deployment Services Deployment Manager クライアント・ソフトウェアへのアクセス権限がないユーザーに特に便利です。

フォルダー。 使用可能なフォルダーが表示され、階層をナビゲートできます。保存先のフォルダーを選択します。フォルダー・リストの上にあるアイコンを使用して、新規フォルダーを作成することや、既存フォルダーの名前を変更することもできます。

ラベルでフィルタリング。 選択したラベルを持つオブジェクトにのみ表示を制限します。フォルダーの参照時には、このラベルを持つオブジェクトのみが表示されます。

アプリケーションで使用しているルールをエクスポートするときには、「名前を付けて保存」ドロップダウンまたは「名前を付けてエクスポート」>ドロップダウンも使用できます。ドロップダウンから「単一のルール・セット」または「個々のルール」を選択して、ルールの保存方法を指定します。参照されているルールのみ使用可能な場合は、ドロップダウンが無効になり、ルールが単一のルール・セットとして保存されず (参照されているルールを個別に保存することはできません)。

個々のルールとして保存することを選択すると、保存する各ルールの詳細を入力するためのテーブルが表示されます。

名前。 保存するオブジェクトの固有の名前を入力します。

キーワード。 保存されたオブジェクトを、検索時に短時間で識別するためのキーワードを入力します。例えば、「販売」 および「1-2009」などと指定します。

説明。 保存されたオブジェクトのより詳細な説明を入力します。例えば、「小売販売の分析、2009 年第 1 四半期」などと指定します。

ラベル。 保存された同じオブジェクトに対して複数のバージョンが発生する可能性がある場合は、各バージョンに固有のラベルを付けて識別することができます。既存のラベルを選択するか、新しいラベルを作成します。

ラベルの操作

ラベルは、プロジェクト、モデル、またはルールの特定のバージョンを識別するために使用します。例えば、複数のプロジェクトでルールを共有する場合に、「実動」というラベルが付いたバージョンを使用するように指定できます。このルールの新しいバージョンが作成された場合でも、「実動」ラベルを新しいバージョンに明示的に移動しない限り、プロジェクトで使用されるバージョンは変更されません。

特に実稼働環境では、共有オブジェクトを参照するときに、ラベル付きのバージョンを使用することを強くお勧めします。これを行わないと、最新バージョンに加えられた変更によって、意図しない結果が発生することがあります。ルール内からルールを参照する場合などの特定のケースでは、ラベル付きのバージョンを使用する必要があります。詳しくは、48 ページの『ルールのエクスポートおよび再使用』のトピックを参照してください。

バージョンが指定されていない場合は、最新バージョンが使用されます。つまり、ルールに対して以降に加えられた変更が、自動的に適用されます (「最新」という特殊なラベルが、新しいバージョンが作成されるたびに引き継がれると考えることができます)。

新しいオブジェクトまたはバージョンを保存するたびに、そのバージョンに適用するラベルを選択できます。詳しくは、25 ページの『プロジェクト、モデル、およびルールのオープンと保存』のトピックを参照してください。

ラベル名の制約。 どのような場合に IBM Analytical Decision Management、文字で使用するラベルを定義するか? #、&、および : (コロン) をラベル名に含めることができません。これらの文字をラベルに使用す

ると、そのラベルを使用するオブジェクトに対する参照を、アプリケーションが正しく解釈できない場合があります。例えば、以下のラベル名は無効です。

テスト & 展開

展開:フェーズ1

プロジェクトのロック

プロジェクト、モデル、またはルールを開くと、「ロック」チェック・ボックスが右上隅で使用可能になり、IBM SPSS Collaboration and Deployment Services Repository 内でそのオブジェクトをロックして、他のユーザーが編集できないようにすることができます。

- プロジェクトがロックされている場合、他のユーザーは、そのプロジェクトを開いて表示することはできますが、変更を保存することはできません。
- ロックされたプロジェクトを閉じると (ブラウザを閉じるなど)、ロックが解除され、他のユーザーがオブジェクトを編集できるようになります。
- プロジェクトを閉じずにロック解除することを選択すると (「ロック」チェック・ボックスを選択解除)、変更を保存するよう求めるプロンプトが出されます。保存していない変更は、オブジェクトがロック解除されると失われます。
- プロジェクトをロック解除すると、そのプロジェクトへの変更を続行できます。ただし、プロジェクトをロック解除しているときに、他のユーザーがそのプロジェクトを偶然ロックした場合を除きます (この場合、ユーザーは変更内容を保存できなくなります)。

プロパティ

「プロパティ」ダイアログには、「プロジェクト」タブと「リポジトリ」タブがあります。「プロパティ」ダイアログを表示するには、任意のアプリケーションで「プロパティ」アイコンをクリックします。



図4. 「プロパティ」アイコン

「プロジェクト」タブで、「名義型」フィールドに設定できる最大メンバー数を指定できます。これは「データ」タブで使用されます。「名義型フィールドの最大メンバー」を選択し、値を指定します。デフォルトは **250** です。これは以下のような場合に役立ちます。

1. ユーザーは新規データ・ソースを作成し、データ・スキャンを実行します。
2. 有用なカテゴリー情報を含む複数のフィールドが「タイプなし」としてプロジェクト・データ・モデルに表示されます。これは、データ内の異なる値の数が、「プロパティ」ダイアログに設定されているデフォルトの最大サイズ制限を超えているためです。
3. ユーザーは、「プロパティ」アイコンをクリックして、「名義型フィールドの最大メンバー」の値をユーザーのデータに適したサイズに増やします。
4. 「データ」タブに戻り、「データ・スキャンを更新」をクリックしても、フィールドは「タイプなし」のままになります。これは、更新アクションによって指標は変更されず、ユーザーが行ったすべての変更が保持されるためです。すなわち、ユーザーがデータ・モデルで行った変更はすべて保持される必要があるため、更新操作はフィールドの既存の指標を保持し、値の再計算のみ行います。

5. 値をプロジェクト・データ・モデルに取り込むには、ユーザーは更新を実行する前に、各フィールドの指標を「タイプなし」から「カテゴリー型」に手動で変更する必要があります。影響を受けるフィールドが多数ある場合は、データ・ソースを再作成する方が速い場合があります。

「名義型フィールドの最大メンバー」設定は、ファイル・タイプ *.sav* のデータ・ソースには適用されないことに注意してください。多くの *.sav* ファイルには、ユーザーがカスタマイズしたファイルと値が含まれているため、この設定による影響を受けません。ユーザーは、フィールドを手動で変更する必要があります。

「リポジトリ」タブで、後日にプロジェクト、モデル、またはルール・セットを参照するときに、それらの識別に役立つキーワードと説明を追加できます。

- **キーワード。** 保存されたオブジェクトを、検索時に短時間で識別するためのキーワードを入力します。例えば、「販売」および「Q1-2013」などと指定します。
- **説明。** 保存されたオブジェクトのより詳細な説明を入力します。例えば、「小売販売の分析、2013 年第 1 四半期」などと指定します。

第 4 章 データ・ソースの管理

「データ」タブでは、分析、シミュレーションおよびテスト、スコアリング、その他の用途のためにデータ・ソースを定義することができます。データ・セットは、管理者が構成するか、またはユーザーが独自に追加できます。

タブは以下の 3 つのメイン領域に分かれています。

- **プロジェクト・データ・モデル**。アプリケーションで必要とするフィールドを定義します。その他のデータ・セットはすべて、このソースに相対的にマップされます。「データ・ソース・フィールド」セクションには、プロジェクトで使用する入力フィールドとタイプがリストされます。必要に応じて、「追加フィールド」セクションを使用して、別のデータ・ソースから式やモデル出力を追加してプロジェクト・データ・モデルを拡張することができます。
- **プロジェクト・データ・ソース**。現在のプロジェクト、モデル、またはルールで使用するために保存されているデータ・ソースがリストされます。
- **マイ・データ・ソース**。自分が定義または保存したデータ・ソースがリストされます。このリストのデータ・セットは異なるプロジェクトとの間でコピーできます。これにより、複数のアプリケーション間でデータを共有することが可能になります。

プロジェクト・データ・モデル

プロジェクト・データ・モデルは、プロジェクトで使用する入力フィールドおよびタイプがリストされたテンプレートのような役割を持っています。

データ・ソース。アプリケーション用に選択されたデータ・ソースを表示します。

キー。アプリケーション用に選択されたデータ・ソースがアプリケーション・ビュー・データ・ソースで、アプリケーション・ビュー・キーがデータ・ソース・テーブルで定義されている場合、このドロップダウンは選択元となるキーのリストを提供します。アプリケーション・ビュー・キーは、IBM SPSS Collaboration and Deployment Services Deployment Manager で作成された IBM SPSS Collaboration and Deployment Services Enterprise View の一部です。アプリケーション・ビュー・テーブル用に定義されたキーがない場合、または、アプリケーション用に選択されたデータ・ソースがアプリケーション・ビュー・データ・ソースでない場合は、「キー」ドロップダウンでは、代わりにデータ・ソースのすべての運用中データ・フィールドのリストが選択項目となります。

選択したキー・フィールドは、「追加フィールド」セクションで定義されたすべての追加データ・ソース・テーブルへのマップに使用されます。実行時、キーはプライマリー・プロジェクト・データ・モデル・テーブルとすべてのセカンダリー・テーブルを結合するために使用されます。プロジェクト・データ・モデルのキー・フィールドは、固有であることが前提です (例えば、顧客 ID、クレーム ID など)。各セカンダリー・テーブルには、一致する列がなければなりません。

キー・フィールドは、データ・ソース用に定義された追加テーブルがない場合はオプションです。

運用中。このボックス内のチェック・マークは、フィールド内のデータがルールと式、および予測に使用できることを意味します。このフィールドを分析のターゲット・フィールドとして使用する場合は、チェック・ボックスからチェック・マークを外します。例えば、ダイレクト・メールによるキャンペーンに対する顧客の回答の度合いを調べるクエリーを実行する場合は、回答値を前もって知ることはできないため、結果フィールドは分析になります。

注: IBM SPSS Collaboration and Deployment Services Enterprise View データ・ソースをプロジェクト・データ・モデルとして選択した場合は、この列のボックスは無効になります。

フィールド名。 データ・ソースにリストされている各フィールドの名前を、ストレージ・タイプを識別するアイコンとともに表示します。

測定。 データ・ソース内の各フィールドの測定タイプを表示します。測定タイプを変更した場合は、データ・スキャンをリフレッシュして、データ互換性を確保する必要があります。詳しくは、32 ページの『測定の尺度』のトピックを参照してください。

値。 データ・ソース内の各フィールドの値をリストします。例えば、範囲内の最大値および最小値です。フラグ・フィールドでは、マウス・ポインターを値の上でホールドすると、そのフィールドに定義されている「true」または「false」の値を示すツールチップが表示されます。値をクリックして編集することもできます。

追加フィールドの追加/編集。 このセクションでは、必要に応じて、別のデータ・ソースから式またはモデル出力を追加して、プロジェクト・データ・モデルを拡張することができます。これを行うには、**追加フィールドの追加/編集**

プロジェクト・データ・ソース

プロジェクト・データ・ソースには、管理者が事前定義したデータ・ソースや、ユーザーが追加したデータ・ソースを含めることができます。オプションで、管理者は、1 つ以上のデータ・ソースをロックして、ユーザーがそれらを変更または削除するのを防止したり、すべてのデータ・オプションをロックして、ユーザーが新しいデータ・ソースを作成できないようにしたりできます。

データ・ソースの入力がプロジェクト・データ・モデルの入力と直接一致しない場合は、前者を後者にマップして不一致を修正できます。例えば、プロジェクト・データ・モデルに「はい」および「いいえ」(測定の尺度・フラグ) の値を持つ「購入」という名前のフィールドが必要な場合、使用するデータ・ソースに同等のフィールドが存在し、それらを適宜マップできなければなりません。

名前。 データ・ソース名を表示し、そのファイル・タイプを識別するアイコンを示します。

レコード数。 この列のアイコンをクリックして、テーブル内の任意のデータ・ソースのレコード数を表示します。数はアイコンのとなりに表示されます。

プレビュー。 この列のアイコンをクリックして、ソースに含まれるデータのサンプルをプレビューします。詳しくは、31 ページの『データのプレビュー』のトピックを参照してください。

概要。 この列のアイコンをクリックして、データ・ソースの概要を表示します。詳しくは、32 ページの『データ概要』のトピックを参照してください。

互換。 データ・ソースがプロジェクト・データ・モデルとして使用されているというメモを表示するか、またはデータ・ソースとプロジェクト・データ・モデルの互換性の度合いを示す、緑、オレンジ、赤のボールを表示します。

- 緑のボールは、データ・ソースとプロジェクト・データ・モデルのデータ・ソースの間に、運用上の互換性があることを示します。運用上の互換性があるデータ・ソースとは、プロジェクト・データ・モデルのすべての運用中フィールドを含み、さらに追加フィールドを持つことができるデータ・ソースのことです。このデータ・ソースは、ルール、スコアリング、シミュレーション、およびテストの運用に適しています。

- オレンジのボールは、プロジェクト・データ・モデルとの互換性がある、同じ名前とタイプのフィールドが少なくとも 1 つデータ・ソースにあることを示します。このデータ・ソースは追加フィールドを持つこともでき、モデルの作成および評価に適しています。
- 赤のボールは、データ・ソースがプロジェクト・データ・モデルとは非互換であり、アプリケーションで使用する前にフィールドのマッピングが必要であることを示します。非互換のデータ・ソースとは、対応するプロジェクト・データ・モデルのタイプとは非互換のタイプを持つフィールドが少なくとも 1 つあるデータ・ソースのことです。

フィールドのマッピング。このオプションを使用すると、データ・ソース・フィールドをプロジェクト・データ・モデルのデータ・ソース・フィールドと比較し、互換性のあるフィールドをマッピングまたはマッピング解除することにより、プロジェクト・データ・モデルが必要とするデータ・ソース・フィールドに一致させることができます。詳しくは、43 ページの『フィールドのマッピング』のトピックを参照してください。

コピー。データ・ソースを「マイ・データ・ソース」領域にコピーできます。

マイ・データ・ソース

このリストのデータ・ソースはあなたのユーザー・アカウントとともに保存されているため、あなたはログインするといつでもこれらを使用できます。また、あなたが開いた任意のプロジェクト、モデル、またはルールにコピーすることができます (ただし、そのための権限が管理者によって付与されていることが前提です)。タブのこの部分にあるフィールドは、「プロジェクト・データ・ソース」領域のフィールドと同じように機能します。ただし、「互換」列はありません。

「コピー」列では、データ・ソースを「プロジェクト・データ・ソース」領域にコピーできます。

データ・ソースの操作

- 新しいデータ・ソースを「プロジェクト・データ・ソース」または「マイ・データ・ソース」のいずれかのリストに追加するには、「データ・ソースの追加」を選択します。詳しくは、34 ページの『新規データ・ソースの作成』のトピックを参照してください。
- 「プロジェクト・データ・モデル」内のデータ・ソースのフィールドの測定の尺度またはタイプを変更するには、該当するレベルを選択します。詳しくは、32 ページの『測定の尺度』のトピックを参照してください。
- 「プロジェクト・データ・ソース」リストとの間でデータ・ソースをコピーするには、「コピー」列内で該当する矢印をクリックします。
- データ・ソースのフィールド名をプロジェクト・データ・モデルにマッピングするには、「互換」列の下の該当するリンクをクリックします (一度フィールド名がマッピングされると、リンクは表示されなくなります)。詳しくは、43 ページの『フィールドのマッピング』のトピックを参照してください。
- データ・ソースをプレビューするには、「プレビュー」アイコンをクリックします。詳しくは、『データのプレビュー』のトピックを参照してください。
- 追加フィールド (別のデータ・ソースの式またはモデル出力) を追加してプロジェクト・データ・モデルを拡張するには、**追加フィールドの追加/編集** をクリックします。詳しくは、42 ページの『追加フィールドの追加』のトピックを参照してください。

データのプレビュー

正しいデータ・ソースを選択したことを確認するために、「プレビュー」アイコンをクリックして、ソースに含まれているデータのサンプルを表示します。プレビューには、「データ」タブの「追加フィールド」セクションにリストされたすべての追加フィールドが含まれます。

データ概要

データ概要機能は、すべてのアプリケーションの「データ」タブ、IBM SPSS Modeler Advantage の「モデル作成」タブと「スコア」タブ、および「ただちにスコアリング」機能が含まれているアプリケーションの「展開」タブで使用可能です。

1. データ・ソースの概要を実行するには、アプリケーション全体で使用可能な「データ概要」アイコンをクリックします。



図5. 「データ概要」アイコン

2. 「データ概要」ダイアログが表示されます。必要に応じて、「概要のオプション」セクションにあるドロップダウンからオーバーレイ・フィールドを選択します。

次に、概要の実行後に、主選択フィールドのみの結果を表示するタブ、またはこのドロップダウンで指定したフィールドに主選択フィールドをオーバーレイするタブが使用可能になります。例えば、結果で、顧客の「年齢」に関する情報を表示してから、「性別」などの別のフィールドでそのフィールドをオーバーレイできます。

3. 必要に応じて、「概要のオプション」セクションの「データ区分データを使用」を選択します。このオプションは、「モデル作成」タブの「オプション設定」でオプション「自動的にデータを区分して、検証およびテストのためのデータ・ソース作成時にモデル評価を実行できるようにする」を選択した場合に使用可能になります。このオプションは、モデルを学習およびテストするために、データを個別のサブセットまたはサンプルに分割します。あるサブセットでモデルを作成し、別のサブセットでそのモデルをテストすることで、そのモデルが他のデータ・セットにどのように一般化されるのかを把握できます。
4. 概要に含めるデータ・フィールドを選択し、「概要の実行」をクリックします。「データ」タブの「追加フィールド」セクションにリストされたすべての追加フィールドを含め、データ・ソースで使用可能なすべてのフィールドがリストされます。デフォルトでは、すべてのモデル入力フィールドおよび対象が選択されています。

データ概要が実行され、結果が表示されます。情報をソートしたり、表示する列を選択したりできます。

5. 任意のフィールドをクリックすると、その詳細が表示されます。選択した各フィールドの新しい結果タブが開き、そこでグラフおよびテーブルを表示したり、使用可能な場合はオーバーレイ・フィールドを選択したりできます。

測定の尺度

測定の尺度は、各データ・フィールドで表される情報を特徴付けるのに役立ちます。また、特定のフィールドがルール、モデル作成、または他のアプリケーションでどのように使用されているのかを判別できます。測定の尺度は、データ・ソースのプロジェクト・データ・モデルで指定できます。例えば、値 0 および 1 を持つ整数フィールドの測定の尺度を「フラグ」に設定して、0 = 偽、1 = 真 と示すことができます。あるいは、使用する入力フィールドを指定する際に、「データ・ソース・エディター」ダイアログでレベルを変更できます。詳しくは、42 ページの『入力フィールドの選択』のトピックを参照してください。

以下の測定の尺度を使用できます。

- **連続。** 0 から 100、または 0.75 から 1.25 の範囲など、数値を記述するために使用されます。連続値は、整数、実数、または日付/時刻にすることができます。
- **カテゴリー。** 異なる値の正確な数が不明である場合に、文字列の値に使用されます。これは**インスタンス化されていないデータ型**です。つまり、データの保管および使用に関して考えられるすべての情報がまだ不明であることを意味します。データが読み取られると、「プロジェクトのプロパティ」ダイアログ・ボックスで指定された名義型フィールドの最大メンバー数に応じて、測定の尺度は、**フラグ型**、**名義型**、または**データ型不明**になります。
- **フラグ。** 何らかの特徴の有無を示す 2 つの異なる値を持つデータに使用されます。例えば、true と false、Yes と No、0 と 1 などです。IBM Analytical Decision Management では、カテゴリー・リスト内の最初の値は「false」の値として使用され、2 番目の値は「true」の値として使用されます。これらは、「データ」タブに表示された値の上にカーソルを合わせると、ツール・チップとして表示されます。データは、テキスト、整数、実数、日付、時刻、またはタイム・スタンプとして表示できます。
- **公称。** small/medium/large など、それぞれがセットのメンバーとして処理される、複数の異なる値を持つデータを記述するために使用されます。公称データには、任意のストレージ (数値、文字列、または日付/時刻) を使用できます。測定の尺度を**名義型**にしても、自動的に値が文字列ストレージに変わることはないことに注意してください。名義型フィールドに許可されている最大メンバー数の設定については、27 ページの『プロパティ』を参照してください。
- **順序。** 固有の順序を持つ、複数の異なる値を持つデータを記述するために使用されます。例えば、給与カテゴリーや満足度ランキングなどは**順序データ**として入力できます。この順番は、データ要素の自然ソート順によって定義されます。例えば、整数セットのデフォルト・ソート順は 1, 3, 5 ですが、文字列セットの順番は HIGH, LOW, NORMAL (アルファベット昇順) です。順序付き測定の尺度を使用して、視覚化、モデル作成、および順序データを **DISTINCT** 型として認識する他のアプリケーション (IBM SPSS Statistics など) へのエクスポートを行うために、順序データとしてカテゴリー・データのセットを定義することができます。名義フィールドを使用できる場所ではどこでも**順序型フィールド**を使用できます。また、任意のストレージ型 (実数、整数、文字列、日付、時刻など) のフィールドも**順序付き**として定義できます。
- **型不明。** 上記のいずれの型とも一致しないデータ、単一の値を含むフィールド、または定義された最大数を超えるメンバーがセットに含まれている公称データに使用されます。これは、「型不明」を使用しなければ、測定の尺度が多数のメンバーを含むセット (お客様番号など) になる可能性のある場合にも便利です。フィールドに「**データ型不明**」を選択すると、役割が自動的に「なし」に設定され、「**レコード ID**」が唯一の代替となります。セットのデフォルト最大サイズは 250 の固有値です。この数値は、ツールバー・アイコンからアクセスできる「プロジェクトのプロパティ」ダイアログ・ボックスで調整または無効化することができます。

データ・フィールド値の編集

データ・テーブルの「**値**」列には、データ・ソース内の各フィールドの値がリストされます。例えば、範囲内の最大値および最小値です。フラグ・フィールドでは、マウス・ポインターを値の上でホールドすると、そのフィールドに定義されている「true」または「false」の値を示すツールチップが表示されます。値をクリックして編集することもできます。ダイアログ・ボックスが開きます。その内容は、選択したデータ・フィールドのタイプによって異なります。以下に例を示します。

- フィールド・タイプが「**連続**」の場合、上限値と下限値を入力できます。
- フィールド・タイプが「**フラグ**」の場合、真の値と偽の値を交換できます。
- フィールド・タイプが「**公称**」の場合、値を追加、削除、並べ替え、および名前変更できます。

新規データ・ソースの作成

1. データ・ソース・エディターにアクセスするには、「プロジェクト・データ・ソース」領域または「マイ・データ・ソース」領域いずれかで「**データ・ソースの追加**」を選択します。
2. データ・ソースの名前を入力します。
3. ソース・タイプを選択します。使用可能なデータベースのリストは、IBM SPSS Modeler マシンで使用可能なデータベースの一覧であることに注意してください。
 - 「**ファイル**」。テキスト (*.txt、*.csv など)、Excel、IBM SPSS Statistics で使用される *.sav および *.zsav 形式、SAS ファイルなど、多くのファイル・タイプがサポートされます。詳しくは、『ファイル・ソース』のトピックを参照してください。
 - 「**データベース**」。Microsoft SQL Server、DB2、Oracle など、ODBC をサポートする任意のデータベース。詳しくは、37 ページの『データベース・ソース』のトピックを参照してください。
 - 「**Enterprise View**」。IBM SPSS Collaboration and Deployment Services で定義されているとおりに IBM SPSS Collaboration and Deployment Services Enterprise View への接続を作成します。詳しくは、38 ページの『「Enterprise View」ソース』のトピックを参照してください。
 - **データ・ビュー**。IBM Operational Decision Management ルールとともに使用される接続を作成します。階層データ・モデルを使用する複雑なルールについては、パフォーマンス低下を招く恐れがあるため管理者がそれらのルールの使用を制限している場合があります。詳しくは、38 ページの『データ・ビュー・ソース』のトピックを参照してください。
 - 「**IBM Cognos BI サーバー**」。IBM Cognos BI サーバーをサポートする任意のデータベース。詳しくは、39 ページの『IBM Cognos BI ソース』のトピックを参照してください。
 - **Analytic Server**。新規データ・ソースを作成するときは、このデータ・ソース・タイプを選択できません。IBM SPSS Modeler 16 では、IBM SPSS Analytic Server が提供され、Hadoop Distributed File System (HDFS) をデータ・ソースとして使用できるようになります。IBM SPSS Modeler ユーザーは、HDFS Analytic Server ソース・ノードを使用してストリームを作成して Hadoop ビッグ・データ・システムに基づいてモデルを構築できます。IBM Analytical Decision Management ユーザーは、これらのモデル・ストリームを IBM SPSS Modeler Advantage で開いて表示できません (Analytic Server データ・ソースを使用して新規プロジェクトを作成することはできません)。
4. ソース・タイプに適切なオプションを指定します。

「**レコード数**」アイコンをクリックして、現在のデータ・ソース内のレコード数を表示できます。あるいは、「**プレビュー**」をクリックして、データ・ソースの最初の 10 件のレコードを表示できます。

ファイル・ソース

アプリケーションでは、いくつかのデータ・ファイル・タイプがサポートされます。ファイル名を入力するか、またはファイルを参照すると、アプリケーションによって自動的にタイプが検出され、データ・ソース・エディターが展開されて、そのタイプに関連する追加フィールドが表示されます。ファイル・タイプが正しくない場合、または特定のバージョン (例えば、Microsoft Excel 1997 から 2003 まで) が必要な場合は、タイプを必要なタイプに変更できます。

「**参照**」をクリックしてリポジトリからファイルを選択できます。あるいは、「**ローカル・ファイルのアップロード**」をクリックして、ローカル・ファイル・システムからファイルを選択できます (有効になっている場合)。

ローカル・ファイルをアップロードした場合、そのファイルは IBM SPSS Modeler Server にアップロードされます。「ファイルのアップロード」ダイアログで、アップロードするローカル・ファイルを参照して選択し、次に IBM SPSS Modeler Server アップロード先を参照して選択します。宛先を選択する際には、必要に応じてファイル名を変更することができます。

デフォルトでは、「ローカル・ファイルのアップロード」機能は無効になっています。管理者は、ブラウザー・ベースの IBM SPSS Collaboration and Deployment Services Deployment Manager を使用して、これを有効または無効にしたり、許可される最大ファイル・サイズ制限を増減したりすることができます。

テキスト・ベースのデータ・ファイル

テキスト・ベースのデータ・ソースを選択すると、さらに詳細を入力するように求めるプロンプトが出されます。

「最初の行は列名」。データ・ソースで見出し行として各列の名前が含まれている場合は、これを選択します。

「エンコード」。使用されるテキスト・エンコード方式を指定します。システム・デフォルトまたは UTF-8 を選択できます。

小数点記号。データ内で小数点がどのように表示されるかを指定します。

- **ピリオド (.)**。ピリオド文字が、小数点として使用されます。
- **コンマ (,)**。コンマ文字が、小数点として使用されます。

区切り文字。このコントロール用にリストされているチェック・ボックスを使用して、コンマ (,) など、ファイル内のフィールド境界を定義する文字を指定できます。複数の区切り文字を使用するレコードの場合、「[,|]」のように複数の区切り文字を指定することもできます。デフォルトの区切り文字は、コンマです。

注: コンマが、小数点としても定義されている場合、このデフォルト設定は使用されません。コンマがフィールド区切り文字と小数点の両方である場合、「区切り文字」リストから「その他」を選択してください。その後、手動で入力フィールドにコンマを指定します。

複数の隣接するブランク区切り文字を単一の区切り文字として使用する場合は、「複数のブランク区切りを許容」を選択します。例えば、1 つのデータ値の後に 4 つのスペースが続き、その後にもう 1 つ別のデータ値が続くグループの場合、5 つのフィールドではなく、2 つのフィールドとして処理されます。

拡張オプション

EOL コメント文字。# や ! などの文字を指定して、データ内の注釈部分を示します。これらのいずれかの文字がデータ・ファイルに表示されている場合、次の改行文字よりも前にある文字はすべて無視されます。

「入力フィールドの指定」。各レコードから使用する入力フィールドの数を指定します。

ヘッダー文字のスキップ。最初のレコードの先頭で無視する文字数を指定します。

型を判別するためにスキャンする行数。指定したデータ型に対してスキャンする行の数を指定します。

「先行および後続のスペースを除去」。インポート時に文字列の先頭および後続のスペースを破棄する場合に選択します。左側、右側、または両側からの除去あるいは削除なしを選択できます。

無効な文字。 データ・ソースから無効文字を削除するには、「**破棄**」を選択します。無効文字を指定した記号 (1 文字のみ) で置き換えるには、「**置換する文字列**」を選択します。無効文字は、ヌル文字、または指定したエンコード方式に存在しないすべての文字です。

引用符。 ドロップダウン・リストを使用して、インポートで単一引用符と二重引用符がどのように処理されるかを指定できます。すべての引用符を「**破棄**」する、フィールド値に引用符を含めて「**テキストとして含める**」、または引用符のペアをマッチングして削除するように「**ペアで破棄**」を行うことができます。引用符が一致しない場合、エラー・メッセージを受け取ります。「**破棄**」と「**ペアで破棄**」では、両方ともフィールド値が (引用符を含まない) 文字列として保存されます。

「**入力フィールドの指定**」。詳しくは、42 ページの『**入力フィールドの選択**』のトピックを参照してください。

Excel データ・ファイル

Excel データ・ソースを選択すると、以下のような、さらなる詳細を入力するように求めるプロンプトが出されます。

「**最初の行は列名**」。データ・ソースで見出し行として各列の名前が含まれている場合は、これを選択します。

「**名前付き範囲**」。Excel ワークシートで定義されているセルの名前付き範囲を選択できます。名前付き範囲を使用すると、その他のワークシートおよびデータ範囲の設定を適用できなくなり、結果として無効になります。

「**ワークシートの選択**」。名前またはインデックスのいずれかで、インポートするワークシートを指定します。

- 「**名前で指定**」。インポートするワークシートの名前を選択します。
- 「**インデックスで指定**」。インポートするワークシートのインデックス値を指定します。インデックス値は 0 (最初のワークシート) から始まり、1 (2 番目のワークシート) というように続きます。

「**ワークシート上の範囲**」。最初の非ブランク行からデータのインポートを開始するか、明示的なセル範囲を指定してデータをインポートすることができます。

- 「**最初の非ブランク行から範囲を開始する**」。最初の非ブランク・セルを検出し、それをデータ範囲の左上隅として使用します。
- 「**セル範囲の明示**」。行と列で明示的な範囲を指定できます。例えば、Excel 範囲 A1:D5 を指定するには、最初のフィールドに A1 と入力し、2 番目のフィールドに D5 と入力します (あるいは、R1C1 と R5C4 を入力します)。ブランク行も含め、指定した範囲内のすべての行が返されます。

「**ブランク行で**」。ブランク行が検出されると、その行をスキップして無視するかどうかを選択できます。または、「**ブランク行を返す**」を選択して、ワークシートの末尾まですべてのデータ (ブランク行を含む) を読み取り続けることができます。

「**入力フィールドの指定**」。詳しくは、42 ページの『**入力フィールドの選択**』のトピックを参照してください。

IBM SPSS Statistics データ・ファイル

IBM SPSS Statistics データ・ソース (.sav または .zsav ファイル) を選択すると、さらに詳細を入力するように求めるプロンプトが出されます。ファイルがパスワードで保護されている場合は、パスワードを入力するように求めるプロンプトも出されます。

「変数名」。IBM SPSS Statistics *.sav* ファイルまたは *.zsav* ファイルからのインポート時に変数名およびラベルを処理する方法を選択します。

- 「名前およびラベルを読み取る」。変数名とラベルの両方を読み取る場合に選択します。これはデフォルト・オプションです。ラベルは、グラフ、モデル・ブラウザー、およびその他のタイプの出力に表示されることがあります。
- 「ラベルを名前として読み取る」。短いフィールド名ではなく、IBM SPSS Statistics *.sav* ファイルの記述変数ラベルを読み取り、そのラベルを変数名として使用する場合に選択します。

「値」。IBM SPSS Statistics *.sav* ファイルまたは *.zsav* ファイルからのインポート時に値およびラベルを処理する方法を選択します。

- 「データおよびラベルを読み取る」。実際の値と値ラベルの両方を読み取る場合に選択します。これはデフォルト・オプションです。
- 「ラベルをデータとして読み取る」。値を表すために使用される数値コードまたはシンボリック・コードではなく、*.sav* ファイルまたは *.zsav* ファイルの値ラベルを使用する場合に選択します。例えば、男性と女性を実際に表す値 1 と 2 が含まれた性別フィールドを含むデータでこのオプションを選択すると、フィールドが文字列に変換されて、男性と女性が実際の値としてインポートされます。

このオプションを選択する前に、IBM SPSS Statistics データにおける欠損値を考慮することが重要です。例えば、数値フィールドが欠損値 (0 = 回答なし、-99 = 不明) にのみラベルを使用している場合、上記オプションを選択すると、値ラベル「回答なし」および「不明」のみがインポートされ、フィールドが文字列に変換されます。そのような場合は、値自体をインポートする必要があります。

「入力フィールドの指定」。詳しくは、42 ページの『入力フィールドの選択』のトピックを参照してください。

データベース・ソース

データベース・ソースを選択すると、以下のような、さらなる詳細を入力するように求めるプロンプトが出されます。

「テスト接続」。データベースのユーザー名とパスワードを入力し、クリックしてデータベースにアクセスできることを確認します。

「テーブル名」。データの使用元のテーブルを選択します。

「先行および後続のスペースを除去」。インポート時に文字列の先頭および後続のスペースを破棄する場合に選択します。左側、右側、または両側からの除去あるいは削除なしを選択できます。

テーブル名および列名を引用符で囲む。データベースにクエリーが送信されるときに、テーブルおよび列の名前を引用符で囲むかどうかを指定します (スペースや句読点が含まれる場合など)。

- 「必要な場合」オプションを使用すると、非標準文字が含まれる場合にのみ、テーブルおよびフィールドの名前が引用符で囲まれます。非標準文字には、非 ASCII 文字やスペース文字に加え、ピリオド (.) 以外のすべての非英数字が含まれます。
- テーブルやフィールドの名前を引用符で囲まない場合は、「行わない」を選択します。
- すべてのテーブル名およびフィールド名を引用符で囲む場合は、「常時」を選択します。

「入力フィールドの指定」。詳しくは、42 ページの『入力フィールドの選択』のトピックを参照してください。

「Enterprise View」ソース

「Enterprise View」ソースを選択すると、以下のような、さらなる詳細を入力するように求めるプロンプトが出されます。

「アプリケーション・ビュー」。ドロップダウンリストに、選択したアプリケーション・ビューが表示されます (存在する場合)。現在のセッションで他のアプリケーション・ビューに対して接続を行った場合、ドロップダウン・リストにはこれらのアプリケーション・ビューも表示されます。「参照」をクリックして、リポジトリ内の他のアプリケーション・ビューを検索します。

「テーブル」。データの使用元のアプリケーション・ビュー・テーブルを選択します。

「データ・プロバイダー」。ドロップダウン・リストには、選択したアプリケーション・ビューを参照するデータ・プロバイダー定義の名前が表示されます。データ・プロバイダー定義を選択するか、「-データ・ソースなし-」オプションを選択します。「-データ・ソースなし-」オプションを選択すると、データ・ソースは実質的にアプリケーションで使用不可になり、テスト、シミュレーションなどで使用できなくなります。

「環境」。環境値 (IBM SPSS Collaboration and Deployment Services で設定) は、どのデータ・プロバイダー定義が使用可能であるかを決定します。

「入力フィールドの指定」。詳しくは、42 ページの『入力フィールドの選択』のトピックを参照してください。

IBM SPSS Collaboration and Deployment Services Enterprise View およびこのデータを構成する詳細な方法については、IBM SPSS Collaboration and Deployment Services の資料を参照してください。今後のリリースでは、IBM SPSS Collaboration and Deployment Services Enterprise View は非推奨になることに注意してください。

データ・ビュー・ソース

データ・ビュー・ソースは、IBM Operational Decision Management ルールとともに使用される接続を作成します。階層データ・モデルを使用する複雑なルールについては、パフォーマンス低下を招く恐れがあるため管理者がそれらのルールの使用を制限している場合があります。詳細については、79 ページの『データ・オプションのロック』を参照してください。

IBM SPSS Collaboration and Deployment Services 分析データ・ビューで定義されたデータをストリームに組み込むように、データ・ビュー・ノードを IBM SPSS Modeler で構成することができます。分析データ・ビューは、予測モデルおよびビジネス・ルールで使用されるエンティティを記述するデータにアクセスするための構造を定義します。ビューにより、分析のための物理データ・ソースにデータ構造が関連付けられます。

詳しくは、IBM SPSS Modeler の資料、および IBM SPSS Modeler ヘルプの『入力ノード、プロセス・ノード、出力ノード』のセクションを参照してください。

データ・ビュー・ソースを選択すると、以下のような、さらなる詳細を入力するように求めるプロンプトが出されます。

分析データ・ソース。 分析データ・ソースをリポジトリから選択します。

テーブル名。 データの使用元のテーブルを選択します。これは、分析データ・ビューのデータ・モデルからのテーブルです。データ・モデルの各テーブルは、予測分析プロセスに関する概念、すなわちエンティティ

ィーを表します。テーブルのフィールドは、テーブルで表されるエンティティの属性に対応します。例えば、顧客注文を分析している場合、データ・モデルには顧客のテーブルと注文のテーブルが含まれている可能性があります。顧客テーブルには顧客 ID、年齢、性別、配偶者の有無、および居住国のフィールドが存在する可能性があります。また、注文テーブルには注文 ID、注文の品目数、総費用、および注文を行った顧客の ID のフィールドが存在する可能性があります。顧客 ID フィールドを使用して、顧客テーブルにある顧客を注文テーブル内のそれぞれの注文に関連付けることができます。

データへのアクセス・プラン。 データの使用元のデータ・アクセス・プランを選択します。データ・アクセス・プランは、分析データ・ビューのデータ・モデル・テーブルを物理データ・ソースに関連付けます。通常、分析データ・ビューには複数のデータ・アクセス・プランが含まれています。使用中のデータ・アクセス・プランを変更すると、プロジェクトで使用されるデータが変更されます。例えば、分析データ・ビューに、モデルの学習のためのデータ・アクセス・プランとモデルのテストのためのデータ・アクセス・プランが含まれている場合、使用中のデータ・アクセス・プランを変更することにより、データの学習からデータのテストに切り替えることができます。

オプションの属性。 必要に応じて、オプションの属性を選択してデータ・ソースに追加します。

「入力フィールドの指定」。詳しくは、42 ページの『入力フィールドの選択』のトピックを参照してください。

IBM Cognos BI ソース

IBM Analytical Decision Management で使用するために、IBM Cognos BI データベースのデータまたは単一リスト・レポートをインポートできます。

注: IBM Analytical Decision Management で使用している IBM SPSS Modeler Server で IBM Cognos BI サーバーがサポートされない場合は、このオプションは使用できません。

「**IBM Cognos BI サーバー**」を選択し、「**接続**」をクリックすると、サーバー接続詳細を入力するように求めるプロンプトが出されます。

1. データのインポート元またはエクスポート先となる IBM Cognos サーバーの「**サーバーの URL**」を入力します。使用する URL が不明な場合は、Cognos システムの管理者に問い合わせてください。
2. 接続する際の「**モード**」を選択します。特定のユーザー (管理者など) としてログインするには、「**資格情報の設定**」を選択し、IBM Cognos の名前空間、ユーザー名、およびパスワードを入力します。
 - サーバーへのログオンに使用する、IBM Cognos セキュリティ認証プロバイダーの「**名前空間**」を入力します。認証プロバイダーは、ユーザー、グループ、および役割の定義と保守に使用されるほか、認証プロセスの制御に使用されます。
 - サーバーへのログオンに使用する IBM Cognos の「**ユーザー名**」を入力します。
 - 指定したユーザー名に関連付けられている「**パスワード**」を入力します。
3. ユーザー資格情報なしでログインする場合は、名前空間、ユーザー名、およびパスワードの各フィールドを指定できないため、「**匿名接続の使用**」を選択します。

注: 一部のサーバー接続では、匿名接続を使用できません。

Cognos オブジェクト詳細の選択

IBM Cognos サーバーへの接続のセットアップ後に、モードを選択します。「**データ**」モードでは、データのインポート元となるパッケージを参照して選択できます。「**リスト・レポート**」モードでは、データのインポート元となる事前定義のリスト・レポートを参照して選択できます。

パッケージには、IBM Cognos モデルと、そのモデルに関連するすべてのフォルダー、クエリー、レポート、ビュー、ショートカット、URL、およびジョブ定義が含まれています。IBM Cognos モデルは、ビジネス・ルール、データ記述、データ関係、ビジネス・ディメンションと階層、およびその他の管理タスクを定義します。

1. パッケージを選択するには、「データ」モードを選択し、「参照」をクリックします。内容をインポートできるすべての使用可能なパッケージをリストしたダイアログ・ボックスが表示されます。
2. あるいは、単一リスト・レポートを選択するには、「リスト・レポート」モードを選択し、「参照」をクリックします。レポートを選択できるパッケージおよびフォルダーをリストしたダイアログ・ボックスが表示されます。注: インポートできるのは単一リスト・レポートのみです。複数リストはサポートされません。
3. 必要な場合、リストを展開して、必要なパッケージまたはレポートを見つけます。項目を選択し、「OK」をクリックします。パッケージまたはレポートの詳細がデータ・ソース・エディターに表示されます。
4. パッケージの選択後に、インポートするオブジェクトを選択します。例えば、パッケージから使用するパッケージ・オブジェクトのクエリー対象物 (データベース表を表す) または個々のフィールド・クエリー項目 (表の列を表す) を選択します。注: データは UTF-8 形式でなければなりません。
5. 生データではなく、集計データをインポートする場合は、「インポート前にデータを集計」オプションを選択します。集計データは一般に、生データよりもはるかにレコード数が少なくなります。

「**オブジェクトの選択**」。パッケージまたはリスト・レポートから選択できるオブジェクト (名前空間とも呼ぶ) を表示します。1 つ以上のオブジェクトを展開して、インポート可能な入力フィールドを表示します。オブジェクトを選択すると、自動的にその入力フィールドがすべて選択されます。

「**パラメーターの確認**」。パラメーターが定義済みの場合、データ・オブジェクトまたはレポートのいずれかに対して、そのオブジェクトまたはレポートのインポート前にこれらのパラメーターに値を指定できます。この表には、フィルターおよび計算に必要なパラメーターのリストが含まれています。これらには、デフォルト値がある場合があります。1 つ以上のパラメーターにデフォルト値がない場合、警告が表示されます。「値」列に、パラメーターを適切に設定します。パラメーターを使用して、データをインポートする前に調整を行うことができます (例えば、パラメーター化された計算を実行するためなど)。

「**フィルターの確認**」。パッケージにフィルターが定義されている場合、1 つ以上のフィルターをインポートできます。インポートしたフィルターがインポートしたデータに関連付けられている場合、データのインポート前にそのフィルターが適用されます。フィルターは、パラメーター化してもしなくても構いません。

「**入力フィールドの確認**」。処理のために IBM Analytical Decision Management にインポートされる入力フィールド (データベース・オブジェクトとも呼ぶ) をリストします。特定のフィールドが不要な場合は、そのフィールド名の隣にあるチェック・ボックスを選択解除します。

Cognos オブジェクト・アイコン: 次の表に示されているように、Cognos BI データベースからインポートできるさまざまなタイプのオブジェクトは、異なるアイコンによって表されます。

表2. Cognos オブジェクト・アイコン:













アイコン	オブジェクト
	パッケージ
	名前空間

表 2. Cognos オブジェクト・アイコン (続き):

アイコン	オブジェクト
	クエリー・サブジェクト
	クエリー項目
	数値データ・ディメンション
	測定
	ディメンション
	レベルの階層
	レベル
	フィルター
	レポート
	単体の計算

Analytic Server ソース

新しいデータ・ソースの作成時に、Analytic Server データ・ソースを選択することはできません。Analytic Server データ・ソースに対して IBM SPSS Modeler で作成されたモデルを開くために IBM SPSS Modeler Advantage を使用するとき、このデータ・ソースは表示専用です。

IBM SPSS Modeler 16 では、IBM SPSS Analytic Server が提供され、Hadoop Distributed File System (HDFS) をデータ・ソースとして使用できるようになります。IBM SPSS Modeler ユーザーは、HDFS Analytic Server ソース・ノードを使用してストリームを作成して Hadoop ビッグ・データ・システムに基づいてモデルを構築できます。IBM Analytical Decision Management ユーザーは、これらのモデル・ストリームを IBM SPSS Modeler Advantage で開くことができます (Analytic Server データ・ソースを使用して新規プロジェクトを作成することはできません)。モデル・タイプやターゲットの変更などの特定のアクションは実行できません。

Analytic Server データ・ソースを「データ」タブに表示するときは、次の情報を選択できます。

「データ・ソース名」。IBM SPSS Modeler でデータ・ソースに付けられた名前。この名前は、必要に応じて変更できます。この後に説明する、フィールドの残りの部分は、編集できません。

「Analytic Server ソース」。IBM SPSS Modeler で構成されているデータ・ソースへのパス。

「入力フィールドの指定」。使用する入力フィールド。

追加フィールドの追加

大部分の IBM Analytical Decision Management アプリケーションでは、式またはモデル出力などの追加フィールドをプロジェクト・データ・モデルに追加できます。「データ」タブの「プロジェクト・データ・モデル」の下のにある「追加フィールド」セクションには、データの一部として組み込まれる予定のすべての式またはモデル出力が表示されます。「追加フィールド」セクションは IBM SPSS Rules Management アプリケーションには含まれていないことに注意してください。

- 「プロジェクト・データ・モデル」セクションの画面の右側で、「追加フィールドの追加/編集」

式を追加するには、以下の手順を実行します。

- 「式の追加」アイコンをクリックします。式エディターが表示されます。詳しくは、43 ページの『式エディター』のトピックを参照してください。

モデル出力を追加するには、以下の手順を実行します。

1. 「モデル出力の追加」アイコンをクリックし、モデルを参照して選択します。選択したモデルの「入力および出力」ダイアログが表示されます。
2. モデルの入力フィールドに互換性がある場合、ダイアログの「モデル入力のマップ」セクションは省略されます。モデルの入力フィールドに互換性がない場合、セクションは展開され、使用可能なテーブル入力値を、選択したモデルの入力値にマップする必要があります。
3. 「モデル出力」セクションに、選択したモデルに使用可能なすべてのモデル出力がリストされます。データに追加するモデル出力フィールドを選択します。「モデル出力名」列には、モデルによって生成されたとおりに、出力フィールドの名前が表示されます。

モデルが出力する任意のフィールドを選択できます。モデルの出力フィールドとは、ストリームの出力先となるすべてのフィールドのことですが、これはストリームへの入力にもなるわけではありません（ストリームによって追加されるすべてのフィールド）。入力フィールドにはプロジェクト・データ・モデル内のものが必ずしもすべて含まれるわけではないため、モデル出力フィールドがプロジェクト・データ・モデル・フィールドと同じ名前である可能性があります。そのため、必要に応じて、ここでデータ・フィールドの名前を変更することもできます。

「測定」列には、該当するタイプのアイコンが先頭に付けられてモデル出力の測定タイプが表示され、「値」列にはモデル出力の値が表示されます。これらの列に結果が表示されるフィールドもありますが、結果が表示されないフィールドもあります。通常、測定タイプと値は、実際のモデル出力（予測値や確率など）についてはデータ・スキャンを行わなくとも既知であるため、既知であれば表示されます。ただし、ストリームからの他の出力フィールド（派生フィールドなど）の測定タイプまたは値は、データ・スキャンなしでは不明な場合があります。

4. 「モデル・パラメーター」セクションは、選択したストリームにスコアリング・パラメーターが含まれる場合、ダイアログの下部に含まれます。モデル・パラメーターの値を入力してください。
5. ダイアログでの操作が終了したら、「保存」をクリックします。選択されたすべてのモデル出力は、「データ」タブの「追加フィールド」セクションに追加されます。

入力フィールドの選択

データが大量にある場合は、ソースからインポートされるデータ量を減らすか、または微調整してください。例えば、関心のある領域に関連するフィールドのみを使用します。

1. 「データ・ソース・エディター」ダイアログの下部にある「入力フィールドの指定」をクリックします。ダイアログが展開され、ソース内のすべてのデータ・フィールドが表示されます。また、定義したすべての式も表示されます。詳しくは、43 ページの『式エディター』のトピックを参照してください。

2. 使用する、それぞれのフィールドおよび式を選択します。

データの選択に加えて、必要に応じて、1 つ以上のデータ・フィールドの測定の尺度を変更することができます。詳しくは、32 ページの『測定の尺度』のトピックを参照してください。

式エディター

式マネージャーを使用して、アプリケーションで使用する追加フィールドまたは属性を派生できます。例えば、銀行データを使用している場合は、顧客の収入と顧客が所有しているローン・アカウントの数の比率を示す式を作成できます。式は常に、「連続」測定の数値です。これを変更することはできません。

注: 運用データと同様に、式はアプリケーション全体で使用可能です。例えば、ルールやモデル予測で使用できます。

- 既存の式を表示するには、「データ」タブで「追加フィールドの追加/編集」をクリックしてから、「式の追加」をクリックします。

式の作成

1. 「式エディター」ダイアログで、式の固有の名前を入力します。
2. データ・ソースから使用可能なすべての入力が表示されます。必要な入力および数学記号をダブルクリックするか、「式の方程式」フィールドにドラッグして、式を作成します。
3. 式の使用目的を説明する「説明」を入力します。
4. 「保存」をクリックします。

フィールドのマッピング

新しいデータ・ソースを追加するとき、そのすべてのフィールドをマッピングして、プロジェクト・データ・モデルとの互換性を保証する必要があります。例えば、プロジェクト・データ・モデルに、値「はい」および「いいえ」(測定の尺度・フラグ)が含まれる「購入」というフィールドが必要な場合、使用するすべてのデータ・ソースに互換フィールドが存在しなければなりません。フィールド名が同一でない場合は、適宜マッピングすることができます。入力フィールドとマッピングされる関連フィールドのデータ型は同じでなければならないことに注意してください。

あるいは、プロジェクト・データ・モデルが変更された場合は、多くのデータ・ソースが正常にマッピングされなくなることがあります。その場合、それぞれのデータ・ソースについて、「プロジェクト・データ・ソース」領域の互換性インジケータ・ボールの色が該当する色に変わります。それらをすべて新しいプロジェクト・データ・モデルに再マッピングする必要があります。

1. 「プロジェクト・データ・ソース」領域で、「フィールドのマッピング」をクリックします。ダイアログが開き、システムによって、フィールドのマッピングについて最適な推測が行われます。任意の時点で、「リセット」をクリックしてシステムの最適な推測に戻すことができます。
2. 「使用可能な入力」列でフィールドを選択し、「必須入力フィールド」列でそれに対応するフィールドを選択してから、「マッピング」をクリックして 2 つのフィールドをマッピングします。「使用可能な入力」列のすべてのフィールドがマッピングされるまで、これを繰り返します。
3. 終了したら、「OK」をクリックします。

第 5 章 ルールの操作

ルールは、ビジネス・ロジックまたは予測モデルの出力に基づいて決定を自動化するために使用されます。アプリケーションに応じて、支払履歴を基に顧客を除外したり、リスクなクレームについて専門調査部門に問い合わせたり、予測モデルで特定された解約、購入、または不正を行う傾向が最も高い顧客を対象としたりする場合にルールを使用できます。ルールは自動化を可能にし、モデルは予測を可能にします。この 2 つを結合できることが、予測アプリケーションを他の分析ツールと区別する重要な特性の 1 つとなっています。

ルールは、各アプリケーションのさまざまな場所で、作成し、適用し、保存することができます。あるいは、を使用して共有ルールの作成と編集を行うことができます。

ルールの作成

- アプリケーション内のルールを追加する場所で、「**新規のルールを作成**」アイコンをクリックします。
- あるいは、「**既存のルールを検索**」アイコンをクリックして、現在のアプリケーションで追加または再利用できる既存のルールを参照します。詳しくは、48 ページの『ルールのエクスポートおよび再使用』のトピックを参照してください。
- 共有ルールを作成するには、アプリケーション起動ページから **IBM SPSS Rules Management** を起動します。

セグメント・ルールの定義

セグメント・ルールは、さまざまなアプリケーションで必要に応じてレコードを選択して割り当てるために、アプリケーションのあらゆる場所で使用されます。セグメントは、真または偽の結果を返す 1 つ以上の式で定義されます。例えば、年齢フィールドの値が 18 未満である場合に真を返す単一のセグメントを含んだルールなどが挙げられます。セグメントは、データ内で定義されるフィールドや式、あるいはモデルから返される傾向スコアや予測に基づいて定義できます。複数のディメンション階層（要素と呼ばれます）で構成された「定義」タブがあるアプリケーションでは、他のディメンション・メンバーから選択することもできます。

各セグメントは 1 つ以上の式で定義され、すべての条件が満たされている場合に真の値を返します。例えば、年齢と性別で定義されたセグメントには、両方の条件を満たしたレコードのみが含まれ、それ以外のすべてのレコードに対してはルールから偽が返されます。

セグメント・ルールを作成または編集するには

1. ルールを追加するアプリケーションで、「**新規のルールを作成**」アイコンをクリックします。
2. ルールの名前を指定します。
3. 入力列で、下矢印アイコンをクリックして、フィールド、ルール、またはモデルを必要に応じて指定します。

表 3. 入力列のオプション：






アイコン	使用法
	有効になっている場合は、以前に定義した要素からディメンション・メンバーを選択できます。

表3. 入力列のオプション (続き) :

アイコン	使用法
	現在のデータ・セットからフィールドを選択できます。
	リポジトリから既存のモデルまたはルールを選択できます。
	新規のモデルを作成できます。
	式を作成できます。

4. 使用する機能を選択します。選択可能なオプションについては、以下の表を参照してください。

表4. 選択可能な演算子 :

値の型	選択可能な演算子
タイム・スタンプ、日付、日付/時刻	<ul style="list-style-type: none"> • = (等しい) • != (等しくない) • > (より大) • < (より小) • >= (より大か等しい) • <= (より小か等しい) • BETWEEN • ISNIL • IS NOT NIL
範囲	<ul style="list-style-type: none"> • = (等しい) • != (等しくない) • > (より大) • < (より小) • >= (より大か等しい) • <= (より小か等しい) • BETWEEN • ISNIL • IS NOT NIL • ONE OF • NOT ONE OF
セット	<ul style="list-style-type: none"> • = (等しい) • != (等しくない) • ISNIL • IS NOT NIL • ONE OF • NOT ONE OF

表 4. 選択可能な演算子 (続き):

値の型	選択可能な演算子
フラグ	<ul style="list-style-type: none"> • = (等しい) • != (等しくない) • ISNIL • IS NOT NIL

5. 入力データ型に適した値を指定します。非数値の値には、ドロップダウン・リストに選択肢が示されません。

あるいは、「既存のルールを検索」アイコンをクリックして、既に定義されているルールを参照します。詳しくは、48 ページの『ルールのエクスポートおよび再使用』のトピックを参照してください。

Age BETWEEN 50 and 20 などのセグメント・ルールでは、結果が返されないことに注意してください。これは、式 $x \leq \text{value and value} < y$ と同じであり、true にはなりません。これは、SQL Server の動作と一貫しています。Age BETWEEN 20 and 50 は、正しい形式です。

セグメント・ルールセット

- セグメント・ルールを結合してセットにすることも可能です。セグメント・ルールセットは、個々のセグメントのいずれかが真である場合に真の値を返します。この点は、すべてが真でなければならないセグメント内部の式と異なるので注意してください。
- セグメント・ルールセットは、ルールを使用する任意のアプリケーションからエクスポートできます。また、使用可能な場合は IBM SPSS Rules Management からエクスポートできます。詳しくは、48 ページの『ルールのエクスポートおよび再使用』のトピックを参照してください。

選択ルールの定義

選択ルールとは、それぞれに包含 または除外 属性が割り当てられた、1 つ以上のセグメント・ルールまたはルール・セットのセットです。選択ルールは、モデル作成、スコアリング、またはその他の処理でレコードのサブセットを選択するために、アプリケーションのあらゆる場所で使用されます。

1. ルールを追加するアプリケーションで、「新規のルールを作成」アイコンをクリックします。このアイコンは、選択ルールを適用できる場所で表示されます。
2. ルールの名前を指定し、ルールを包含と除外のどちらに使用するかを選択し (後で行うことも可能)、必要に応じて 1 つ以上の式を追加します。詳しくは、45 ページの『セグメント・ルールの定義』のトピックを参照してください。
3. 「OK」をクリックしてルールを保存します。
4. ルールごとにこの設定を切り替える場合は、「包含/除外」列のアイコンをクリックします。選択ルール・セットが参照されている場合は、この設定を変更できません。同じルール内で、包含と除外を組み合わせることはできません。

あるいは、「既存のルールを検索」アイコンをクリックして、既に定義されているルールを参照します。詳しくは、48 ページの『ルールのエクスポートおよび再使用』のトピックを参照してください。

選択ルールの評価

「選択ルールロジックの表示」アイコンをクリックして、どのように選択が行われるか (論理積か論理和か) を確認します。

- 選択ルールが評価される際に、複数の包含セグメントが指定されている場合は、セグメントのいずれかと一致するレコードが包含されます。例えば、信用格付けまたは保有期間に基づいて顧客を包含することができます。
- 複数の除外セグメントが指定されている場合は、セグメントのいずれかと一致するレコードが除外されます。例えば、年齢または信用格付けに基づいて顧客を除外することができます。
- 競合が発生した場合は、除外ルールが優先されます。例えば、除外セグメントのいずれかに該当した顧客は、適格性が認められるセグメントの有無にかかわらず除外されます。
- IBM SPSS Rules Management で選択ルール・セットを作成する場合は、すべてのルールが同じタイプ (包含 または除外) でなければなりません。同じルール・セット内で、包含ルールと除外ルールを組み合わせることはできません。

ルールのエクスポートおよび再使用

デフォルトでは、アプリケーションで作成されたルールは、そのアプリケーションを使用して保存されます。あるいは、複数のアプリケーションで使用するために、ルールを IBM SPSS Collaboration and Deployment Services Repository にエクスポート、すなわち保存できます。この場合、各アプリケーションは、共有ルールのラベル付けされた特定のバージョンへの参照を保管します。ルールに対する変更は、そのバージョンを使用するすべてのアプリケーションに自動的に適用されます。

ルールのエクスポート

共有するローカルルールが含まれている任意のアプリケーションから、「エクスポート」アイコンをクリックして、そのルールを、必要に応じて再使用できる IBM SPSS Collaboration and Deployment Services Repository に保存します。

- 複数のルールをエクスポートする場合は、すべてのルールを単一のルール・セットでエクスポートするか、個々のルールをエクスポートするかを指定します。個々のルールの場合は、各セグメントが別個のルールとしてエクスポートされます。
- 複数の選択ルールをエクスポートする場合は、「**選択内容とともに保存**」を選択すると、選択ルール・セットがエクスポートされます。この場合、包含または除外のいずれかをエクスポートできますが、両方をエクスポートすることはできません。包含と除外の両方が必要な場合は、別個のルール・セットでそれらをエクスポートする必要があります。「**選択内容とともに保存**」が選択されていない場合は、セグメント・ルールセットがエクスポートされます。
- 集計ルール・エディターからエクスポートする場合は、「**リスク・ポイントと共にエクスポート**」を選択すると、集計ルール・セットが作成されます。このオプションを選択しなかった場合は、セグメント・ルールセットがエクスポートされます。

既存のルールの使用

- 共有ルールを使用する任意のアプリケーションで、「**既存のルールを検索**」アイコンをクリックして、IBM SPSS Collaboration and Deployment Services Repository で、現在のアプリケーションで再使用できる既存のルールを参照します。
- ラベル付けされた特定のバージョンを参照するには、ルール名の隣にある矢印をクリックし、目的のバージョンを選択します。バージョンが選択されていない場合は、最新バージョンが使用されます。(「最新」という特殊なラベルが、新しいバージョンが作成されるたびに引き継がれると考えることができます)。ルールの新しいバージョンを作成すると発生する可能性がある、予期しない結果を避けるために、ラベル付けされた特定のバージョンを使用することをお勧めします。

参照されているルールは、多くの場合、ローカルルールと結合できますが、太字テキスト、およびルール (またはルール・セット) のタイプを示すアイコンによって、ローカルルールとは視覚的に区別されません。

- あるいは、必要に応じて、参照されているルールをローカルルールに変換して、現在のアプリケーションに保存できます。これを行うには、任意の参照されているルールの名前をクリックしてから、「ルールのプレビュー」ダイアログ・ボックスで「単一ルールへの変換」を選択します。

ルールを共有するためのガイドライン

共有ルールを参照して処理する際には、以下のガイドラインが適用されます。

- ルールは、再使用するために任意のアプリケーションからエクスポートできます。ただし、エクスポート後には、IBM SPSS Rules Management で編集できますが、他のどのアプリケーションでも編集できません。IBM SPSS Rules Management は一部のサイトでインストールされていない場合や、一部のユーザーが使用できない場合があることに注意してください。
- 特に実稼働環境では、ラベル付けされた特定のバージョン (*LATEST* 以外) を使用することを強くお勧めします。そうしないと、最新のバージョンに対する継続的な変更により、予期せぬ結果が生じることがあるからです。ルール内からルールを参照する場合などの特定のケースでは、ラベル付きのバージョンを使用する必要があります。
- 単一ルールは、参照によって IBM SPSS Collaboration and Deployment Services Repository から追加するか、「ルールのプレビュー」ダイアログ・ボックスからローカルルールに変換することができます。
- 単一ルールは、任意のルール・エディターで、参照したり、ローカルルールと結合したりできます。ルール・セットとしてエクスポートした場合、単一ルールに対するネストされた参照は保持されます。
- セグメント・ルールセットと選択ルール・セットは、ほとんどの場合、単一ルールと同様に参照したり結合したりできますが、結合されたルール・セットは、まずネストされたルール・セットをすべて単一ルールに変換しなければ、エクスポートできません。そのため、エクスポートされたルール・セットには、ネストされたルールへの参照が含まれることがありますが、ネストされたルール・セットへの参照は含まれません。
- 集計ルール・セットも参照できますが、他のルールと結合することはできず、また他のルールへのネストされた参照を含むこともできません。集計ルール・セットを参照する場合には、参照されているルールは、集計ルール・エディターの内容全体を置き換えます。(あるいは、集計ルール・セットを参照しない場合は、集計エディターは、ローカルルール、参照されているセグメント・ルール、およびセグメント・ルールセットの混合を含むことができます。)
- IBM SPSS Rules Management を使用する場合、エクスポートされるルールの場合と同じ制限が適用されます。ネストされたルールは許可されますが、ネストされたルール・セットは許可されません。これは、IBM SPSS Rules Management が、(他の任意のアプリケーションからのエクスポートと同様に) 共有ルールを作成および編集するように設計されているためです。
- 参照されているルール・セットのすべての制限を回避するために、IBM SPSS Rules Management で作業している場合でも、他の場所で作業している場合でも、ルール・セットは常に別個のローカルルールに変換できます。
- 共有ルールには、外部の IBM Operational Decision Management ルールへの参照を含めることはできません。

注釈の追加

集計または割り当てで、「注釈の追加」アイコンをクリックして、ルール出力に含める注記を入力できるフリー・テキスト・フィールドの列を追加します。例えば、これは、保険請求に問い合わせのフラグが立てられている場合の連絡先をコール・センターのオペレーターに知らせる追加情報である場合が考えられます。

注釈を、ディメンションに指定された「次とともに返す」の値と組み合わせて使用して、追加詳細を提供できます。「次とともに返す」の値は、返されるディメンションまたは決定に関連付けられていますが、注釈は、その決定への入力として起動されたルールを知らせます。

IBM Operational Decision Management からのルールの使用

IBM Analytical Decision Management のアプリケーションでは、IBM Operational Decision Management などのビジネス・ルール管理システムで作成したルールを参照して使用できます。これらの外部ルールは、アプリケーションの任意のポイントで「既存のルールを検索」アイコンをクリックすることにより、他のルールと同様の方法で選択および使用できます。外部ルールはアイコンで識別されますが、それ以外の点では通常の方法で使用できます。

IBM Analytical Decision Management で使用する外部ルールの作成については、「アプリケーション・デザイナーズ・ガイド」を参照してください。

これらのルールのアプリケーションへの挿入は、他のルールと同様の方法で行います。ルール出力が、ルール使用コンテキストと一致している必要があります。例えば、割り当ての場合は、決定に対する入力として、true/false またははい/いいえ (ブール型) の値が必要です。

アプリケーションで使用する任意の外部ルールの詳細 (入力やフィールド、ルールの説明やその他のプロパティなど) を表示するには、他のルールと同様、ルール名前をクリックします。

外部ルールに複数の出力フィールドがある場合は、必要なフィールドを選択できるプロンプトが表示されません。

外部ルールをモデルの代わりに使用する場合は、ターゲット・フィールドが表示されず、すべての構成済み出力がドロップダウン・リストに表示されます。

外部ルール使用のヒント

- *true* または *false* (もしくははい/いいえまたは 0/1) を返すルールは、任意の選択または割り当てにおいて、ルール・エディターで任意のルール入力として使用できる他、特定のディメンション項目に関連付けられた「次とともに返す」フィールドとして使用できます。
- カテゴリー出力 (赤、白、または青 や、北、南、東、西 など) を返すルールは、「定義」タブの「決定のためにモデルを使用する」セクションで、ルール・エディターに対する入力として使用できる他、ディメンション項目に関連付けられた「次とともに返す」フィールドとして使用できます。
- 数値型の出力を返すルールは、「定義」タブの「決定のためにモデルを使用する」セクションで、優先順位付けに対する入力、ルール・エディターに対する入力、またはディメンション項目に関連付けられた「次とともに返す」フィールドとして使用できます。
- 分析データ・ビュー・データ・ソースは、複雑な (階層型の) 外部ルールの使用をサポートします。複雑な外部ルールを参照するには、プロジェクト・データ・モデルがデータ・ビュー・ソースであり、機能が使用可能になっている必要があります (「データ」タブで、管理者は該当する「階層データに基づく ODM 規則の使用の許可」オプションを選択する必要があります)。

- IBM Analytical Decision Management で、外部ルールへの参照を含む共有ルールを作成することはできません。

第 6 章 予測モデルの作成

予測モデルについて

モデルは、過去のデータに存在するパターンに基づいて、将来発生する可能性の高い事象を予測するために使用します。例えば、モデルを使用することで、チャーン行動（契約/解約を繰り返す顧客の流動現象）を取る可能性が最も低い顧客や、特定のオファーに応じる可能性が最も高い顧客を、収入や年齢、加入している組織やメンバーシップなどの特性に基づいて予測できます。

モデルは、ルールと同様の方法で使用できますが、ルールが会社の方針やビジネス・ロジックなどの前提条件に基づくものであるのに対し、モデルは、過去の結果で実際に見られた事象に基づくものであり、他の方法では判明しない可能性のあるパターンを発見できます。ルールが、共通のビジネス・ロジックをアプリケーションに組み込むものであるのに対し、モデルは、洞察と予測能力を付加するものです。

モデルの作成およびスコアリングのためのデータ

モデル作成プロセスでは、以下の 2 つの異なるタイプのデータが使用されます。

- モデルを作成するために、予測対象に関する情報が必要です。例えば、チャーン行動を予測する場合は、過去にチャーン行動を取ったことがある顧客に関する情報が必要です。これは、多くの場合、履歴データまたは分析データと呼ばれ、プロジェクト・データ・モデル内のフィールドの一部またはすべてと、予測対象の結果を記録する追加のフィールドが含まれている必要があります。この追加のフィールドは、モデル作成のターゲットとして使用されます。
- 将来の結果の予測でモデルを使用するために、対象となるグループまたは母集団に関するデータ（潜在的顧客や受け取ったクレームなど）が必要です。これは、多くの場合、運用中データ、またはスコアリング・データと呼ばれます。プロジェクト・データ・モデルは、通常、このデータに基づいています。
- ターゲット・フィールドがデータ・モデルに含まれている場合は、このフィールドの「運用中」列を選択しないようにする必要があります。これは、このフィールドをモデルのスコアリングに使用することはできないためです。モデルは履歴データに基づいて作成されているため、スコアリングの目的は、結果がまだわからない新規データにモデルを適用することです。

モデルの作成

「モデリング」タブでは、データ・ソース、ターゲット・フィールド、およびその他のオプションを指定できます。作成するモデル・タイプも選択できます。

IBM SPSS Modeler でアプリケーション（ストリーム・ファイル）にパラメーターが追加されている場合は、「モデリング」タブと「スコア」タブで「パラメーター」リンクが使用可能になります。「パラメーター」リンクをクリックすると、定義済みのパラメーターとその現在の値が表示され、パラメーターの値を変更して、モデルの作成またはスコアリングに適用できるようになります。

予測モデルの作成

予測モデルを作成するときに指定する必要があるのは、データ・ソースとターゲット・フィールドのみです。データの準備および区分化のためのオプションや、使用するフィールドおよびレコードを選ぶためのオプションを選択することもできます。モデルは、ターゲット・タイプに応じたさまざまな方法を適用し、選択したデータに最も適した特定の手法および結合を自動的に識別することによって生成されます。

モデルをスコアリングすると、予測を含む 1 つ以上のフィールドと、それらの予測に関連付けられた信頼値が返されます。例えば、ターゲット・フィールドが「応答」という名前のフラグ・フィールドである場合、そのフィールドの予測値を含む「XF-応答」という名前のフィールドがモデルによって返されます。

1. データ・ソースを指定します。これには、予測する結果を記録する任意の分析データ・ソースを指定できます。
2. 予測する結果を記録するターゲット・フィールドを指定します。例えば、解約した顧客、過去のオファーに応じたことのある顧客、不正なクレームを行ったことのある顧客をなど示すターゲット・フィールドを指定することが考えられます。

他のすべてのフィールドは、この値の予測に役立つ可能性のある入力として自動的に追加されます。

3. 必要に応じて、オプション設定を指定します。詳しくは、60 ページの『オプションのモデル設定』のトピックを参照してください。
4. 必用な場合は、「データ概要」アイコンをクリックして、現在のモデルの作成に使用するデータの概要を確認します。詳しくは、32 ページの『データ概要』のトピックを参照してください。
5. 「**モデルの作成**」をクリックします。

モデルの作成中は、ブラウザを閉じることや、他のタスクを処理することができます。作成が完了すると、ギャラリーからモデルにアクセスできるようになります。詳しくは、23 ページの『ギャラリー』のトピックを参照してください。モデルの作成に予想以上に時間がかかる場合は、「**停止**」ボタンをクリックして、前に保存したバージョンに戻ることができます。

モデルの作成が完了すると、結果が表示されます。表示内容として、「**結合したモデル結果**」(作成されたばかりの結合モデルのパフォーマンス・グラフ) または「**個々のモデル結果**」(結合モデルを構成する個々のモデルの高水準の詳細) を選択できます。

「個々のモデル結果」の表示時には、各モデルの名前がリンクになっており、ここからモデルに関する情報(ターゲット・フィールドや入力フィールドなど)を含む新しいダイアログを開くことができます。選択したモデルがツリー・モデルの場合は、ツリー・ビューアーも表示されます。

6. オプションで、評価機能やテスト機能を使用して、サンプル・データに対するモデルの動作を確認できます。詳しくは、63 ページの『モデルの評価』のトピックを参照してください。

予測モデル作成の結果

予測モデルが完成すると、結果セクションで「**結合したモデル結果**」と「**個々のモデル結果**」の表示を切り替えることができます。「**結合したモデル結果**」には、作成された結合モデルの分布図および変数の重要度グラフが含まれており、モデルのパフォーマンス、および最終結果を得る際の各種予測値の重要度を簡単に目で見て評価できます。「**個々のモデル結果**」には、結合モデルを構成する個々のモデルに関する情報がすべてリストされ、個々のモデルをクリックすると、その詳細を表示することができます。

分布図

モデルの分布図では、観測値が横軸に表示され、予測値が縦軸またはオーバーレイに表示されます。これにより、各値または反応が予測された頻度と、その予測が正しかった頻度が表示され、モデルの精度を簡単に目で見て評価できます。

- フラグ型対象またはセット型対象の場合は、分布が棒グラフで示されます。この場合、ターゲット・フィールドのカテゴリごとに 1 つのバーが表示されます。各バーの高さは、モデルによって結果が予測された回数を示します。各バーの色分けされた領域は、一致する結果がテスト・データで実際に観測された回数を示します。モデルの精度が妥当であれば、各バーの最大領域が対応する反応と一致するはずですが。

- 範囲型対象の場合は、分布がビン分割散布図で表されます。ここでは、予測値が横軸に作図され、観測値が縦軸に作図されます。点が対角線に沿って集まるのが理想的です。点（結果）が散在しているほど、また線形分布から離れているほど、モデルの精度は低いことになります。

予測の重要度グラフ

多くのユーザーがモデル作成時に重視するのは重要度の高い予測フィールドであり、重要度の低いフィールドは削除するか破棄することを検討します。予測の重要度グラフを使用して、モデルを評価するときに予測値ごとに相対的な重要度を示すことで、この作業に役立たせることができます。これらの値は相対的であるため、表示されるすべての予測値を合計すると 1.0 になります。予測の重要度は、モデルの精度には関係しません。これは、予測値を作成するときに、各予測値の重要度に関係するのみであり、予測値が正確であるかどうかには関係しません。

特に大容量のデータ・セットを使用している場合、予測の重要度の計算はモデルの作成よりもかなり長い時間かかる場合があります。

モデル作成時にデータを自動的に区分化するオプションを選択した場合は、「評価」ページでテスト・データ区分が自動的に選択されます。

区分化を有効にしなかった場合は、モデルを評価するためのデータ・ソースを選択する必要があります。モデルの作成に使用したデータ・ソースとは異なるデータ・ソースを選択することをお勧めします。区分化はデフォルトで有効になるため、その設定を変更しない限り、有効にする必要はありません。

個々のモデル結果

「個々のモデル結果」には、結合モデルを構成する個々のモデルに関する情報が表示されます。各モデルの名前には、リンクが付いています。このリンクをクリックすると、新しいダイアログが開き、そのモデルに関する詳細（ターゲット・フィールドや入力フィールドなど）が表示されます。選択したモデルがツリー・モデルの場合は、特殊なツリー・ビューアーも表示されます。ツリー・ビューアーは対話式であり、以下の機能を備えています。

- ツリーをズームインおよびズームアウトする
- ツリー項目を表、グラフ、または表とグラフの組み合わせで表示する
- ツリーの方法を上から下、左から右、または右から左へと変更する
- ツリー・マップ・ウィンドウ と矢印アイコンを使用して、大きなツリーをより簡単にナビゲートする

予測ルール・モデルの作成

予測ルール・モデルを使用すると、特定の結果と比較して、より可能性が高い、または低いセグメントを識別できます。例えば、チャーン行動（契約/解約を繰り返す顧客の流動現象）を取る可能性が最も低い顧客や、特定のオファーに応じる可能性が最も高い顧客を探すことができます。結果のモデルは、選択ルールのリストに似ており、各セグメントのスコアリング方法を指定する「包含」属性または「除外」属性が含まれます。

予測ルール・モデルは選択ルールに似ていますが、スコアリングの方法が異なります。選択ルールの場合は、除外されたレコードが削除されるため、出力に含まれるレコードが少なくなります。モデルの場合は、レコードが削除されることはなく、各レコードの予測値と傾向値を示す別のフィールドが追加されます。包含されるセグメントのスコアリング結果は、*True*（または *False*、もしくは予測しようとする任意の結果）です。除外されるセグメントのスコアリング結果は *NULL* で、これは、どのルールにも合致しないために残りのレコードとなるすべてのレコードと同じです。

各レコードには、ターゲット応答の可能性を示す傾向スコアも割り当てられるため、オファーに応じる可能性の最も高い顧客を識別することができます。このため、予測ルール・モデルは、メーリング・リストの生成に適しており、コール・センターや販売促進アプリケーションなどのカスタマー・リレーションシップ・マネジメントに広く使用されています。

- 予測ルール・モデルは、カテゴリ・ターゲットに対してのみ使用できます。
- セグメントを表示、変更、および再編成して、スコアリングの対象とするセグメントまたは除外するセグメントを選択できます。例えば、将来のオファーの対象から特定の顧客グループを除外して、他のグループを追加し、これが全体的なヒット率にどのように影響するかを直ちに確認できます。
- 選択ルールと同様、必要に応じて手動でセグメントを追加または編集できます。ただし、通常は、モデルを使用して自動的にセグメントを検出する方が簡単で、結果の微調整に役立つオプションも数多く用意されています。
- ルールは、リストされている順序に従って適用され、最初に一致したルールが特定のレコードの結果を決定します。ルールまたは条件を個別に見ると重複している場合でも、ルールの順序によってあいまいさが解決されます。一致するルールがないレコードは、「残り」ルールに割り当てられます。

予測ルール・モデルを取得するには

1. データ・ソースを指定します。これには、予測する結果を記録する任意の分析ソースを指定できます。
2. 予測する結果を記録するターゲット・フィールドを指定します。
3. 検出する特定の値または応答を指定します。例えば、ターゲット・フィールドの名前が「応答」である場合、*True* または *False* を選択することで、それぞれ応答した、または応答しなかった顧客を検出できます。

例えば、チャーン行動を取ったことがある顧客、過去のオファーに応じたことがある顧客、不正なクレームを行ったことがある顧客などを示すターゲットを指定できます。他のすべてのフィールドは、この値の予測に役立つ可能性のある入力として自動的に追加されます。

4. 必要に応じて、オプション設定を指定します。詳しくは、60 ページの『オプションのモデル設定』のトピックを参照してください。
5. 必要な場合は、「データ概要」アイコンをクリックして、現在のモデルの作成に使用するデータの概要を確認します。詳しくは、32 ページの『データ概要』のトピックを参照してください。
6. 「作成の開始」をクリックします。
7. 指定したターゲットと比較して可能性が高いセグメントと低いセグメントのどちらを検出するかを指定します。例えば、チャーン行動を取る可能性が最も高いグループまたは顧客と、その可能性が最も低いグループまたは顧客のどちらを検出するかを指定し、その結果に従って対象に含めることができます。
8. 検出するセグメントの最大数を指定します。この数値は、通常、処理の単純化とパフォーマンスを考慮して、低く抑えられます。
9. 「モデルの拡張」をクリックして、可能性がより高い、またはより低いグループを識別する 1 つ以上のセグメント・ルールのリストを取得します。
10. 「包含」/「除外」アイコンをクリックして、ルールの使用方法、および該当セグメントのスコアリング方法を指定します。

包含されるセグメントには、ターゲット値 (応答 = *true* など) と一致するスコアが付与されます。除外されるセグメントには、*NULL* というスコアが付与されますが、ファイルからは削除されません。

11. オプションで、必要に応じて手動でセグメントを追加または編集できます。詳しくは、45 ページの『セグメント・ルールの定義』のトピックを参照してください。

12. オプションで、評価機能やテスト機能を使用して、サンプル・データに対するモデルの動作を確認できます。

ヒント:

- ルールを上下に移動するには、上/下矢印を使用します。各レコードは最初にヒットしたルールに基づいてスコアリングされるため、ルールの順序によって結果が変わります。例えば、複数のルールに一致するレコードは、最初に適用されるルールに基づいてスコアリングされます。
- 1 つ以上のオプションを変更して「**モデルの拡張**」を再度クリックすると、追加のルールを検出できます。例えば、追加のセグメントを識別するために、最小セグメントのしきい値や信頼区間を小さくすることができます。
- 「**モデルの拡張**」をクリックしても、テーブルがクリアされることや、モデルが最初から再作成されることはなく、テーブルに追加する別のセグメント・ルールが検出されるだけです。

クラスタリング・モデルの作成

クラスタリング・モデルを使用すると、レコードを特定数のクラスターにカテゴリ分類できます。これは、データ内の自然発生的なグループを識別するために役立ちます。

クラスタリング・モデルは類似するレコードのグループを特定し、レコードが属するグループに従って、レコードにラベルを付けることに重点を置いています。この作業は、各グループとそれらの特性に関する事前の知識を活用せずに実行されます。実際には、検索するグループ数が正確にわからない場合もあります。これが、クラスタリング・モデルと他のマシン学習手法との違いであり、クラスタリング・モデルには、モデルによって予測を行う定義済みの出力フィールドやターゲット・フィールドはありません。クラスタリング・モデルは、モデルの分類性能を判定する外部標準がないため、**教師なし学習モデル**と呼ばれることがよくあります。これらのモデルには、**正** または **誤** という答えはありません。モデルの価値は、データ内で関心のあるグループ構成を取得し、それらのグループ構成について役に立つ説明を提供できるかどうかによって決まります。

クラスタリング手法は、レコード間およびクラスター間の距離の測定に基づいています。レコードは、同じクラスターに属するレコード間の距離を最小にするようにして、クラスターに割り当てられます。

クラスタリング・モデルは、クラスターやセグメントを作成するためによく利用されます。このクラスターやセグメントは、後の分析で入力として使用されます。一般的な例として、マーケット・セグメントがあります。これは、マーケティング担当者がマーケット全体を等質のサブグループに区分するために使用します。各セグメントには、そのセグメントを対象としたマーケティング作業の成功に影響を及ぼす特性があります。データ・マイニングを使用してマーケティング戦略を最適化している場合は、適切なセグメントを識別し、そのセグメント情報を予測モデルで使用することで、多くの場合、モデルを大幅に改善できます。

クラスタリング モデルを取得するには

1. データ・ソースを指定します。これには、予測する結果を記録する任意の分析ソースを指定できます。
2. 必要に応じて、オプション設定を指定します。詳しくは、60 ページの『オプションのモデル設定』のトピックを参照してください。
3. 必用な場合は、「データ概要」アイコンをクリックして、現在のモデルの作成に使用するデータの概要を確認します。詳しくは、32 ページの『データ概要』のトピックを参照してください。
4. 「**クラスターの検索 (Find Clusters)**」をクリックします。
5. オプションで、手動クラスターを追加できます。詳しくは、58 ページの『手動クラスターの使用』のトピックを参照してください。

6. オプションで、評価機能やテスト機能を使用して、サンプル・データに対するモデルの動作を確認できます。

手動クラスタの使用

必要に応じ、クラスタを手動で定義してから、自動クラスタリングを使用して、手動で定義したクラスタでは識別されなかったレコードのクラスタを検出することができます。手動クラスタの作成は、「最初のヒット」に基づくセグメント・ベースの割り当てルール・セットに従って行います。ユーザーが作成したルールによってクラスタが作成され、その残りがクラスタ・モデルの自動作成に使用されます。

手動クラスタは、結果要約グラフには含まれません。

1. 「手動クラスタ」セクションで、「新規のルールを作成」アイコンをクリックし、手動クラスタを定義する 1 つ以上のルールを追加します。
2. 必要に応じ、クラスタのラベル名を変更します。
3. 「手動クラスタのレコード」アイコンをクリックして、手動クラスタに含まれるレコードの総数のレコード・カウントを実行します。定義した各手動クラスタの「カウント」列が更新されます。

アソシエーション・モデルの作成

アソシエーション・モデルを使用すると、一緒に出現する可能性が最も高い項目を予測し、それらの項目間の関係の強度を予測できます。

アソシエーション・ルールは、特定の結論 (特定の製品の購入など) と条件セット (複数の他の製品の購入など) を関連付けます。例えば、以下のルールがあるとします。

```
beer <= cannedveg & frozenmeal (173, 17.0%, 0.84)
```

このルールは、*cannedveg* と *frozenmeal* が一緒に発生すると、*beer* が発生することが多いことを示しています。このルールの信頼性は 84% であり、データの 17%、つまり 173 個のレコードに適用されます。アソシエーション・ルールのアルゴリズムは、Web ノードなどの視覚化技術を使用すれば手動で見つけることができるアソシエーションを自動的に検出します。

より標準的なディジション・ツリー・アルゴリズム (C5.0 および C&R ツリー) よりもアソシエーション・ルール・アルゴリズムのほうが優れている点は、アソシエーションがすべての属性間に存在し得るということです。ディジション・ツリー・アルゴリズムは単一の結論のみに導くルールを構築しますが、アソシエーション・アルゴリズムはそれぞれが別の結論に導く可能性のある多数のルールを見つけようとします。

アソシエーション・アルゴリズムの欠点は、非常に大規模になる可能性のある検索スペース内でパターンを検出しようとするため、ディジション・ツリー・アルゴリズムの実行よりもかなり長い時間を要する可能性があるということです。このアルゴリズムでは、ルールの検出に生成と検定の手法を使用します。この手法では、最初に簡単なルールが生成され、そのルールの妥当性がデータ・セットに対して検証されます。有効なルールは保存され、その後、さまざまな制約に従ってすべてのルールが特殊化されます。特殊化とは、ルールに条件を追加する処理のことです。次に、データに対してこれらの新規ルールが検証され、この繰り返しによって、検出されたルールの中で最善または関心度の高いいくつかのルールが保存されます。通常、ユーザーは、ルールで使用できる前提条件数を制限し、情報理論に基づく各種の手法や効果的なインデックス作成方法を使用して、膨大になる可能性がある検索スペースを縮小します。

アソシエーション・モデルを取得するには

1. データ・ソースを選択します。これには、予測する結果を記録する任意の分析ソースを指定できます。

2. データ形式を指定します。詳しくは、『テーブル形式データとトランザクション・データ』のトピックを参照してください。
3. 必要に応じて、オプション設定を指定します。詳しくは、60 ページの『オプションのモデル設定』のトピックを参照してください。
4. 必用な場合は、「データ概要」アイコンをクリックして、現在のモデルの作成に使用するデータの概要を確認します。詳しくは、32 ページの『データ概要』のトピックを参照してください。
5. 「モデルの作成」をクリックします。
6. オプションで、テスト機能を使用して、サンプル・データに対するモデルの動作を確認できます。アソシエーション・モデル・タイプでは「評価」は使用できません。

他のモデル・タイプでは、別のモデリング手法の中から選択できます。アソシエーション・モデルの場合は、*A priori* モデル・タイプが使用されます。

テーブル形式データとトランザクション・データ

アソシエーション・モデルの作成時には、データ形式として「テーブル形式」または「トランザクション」を使用できます。デフォルトではテーブル形式データが選択されています。「トランザクション」を選択する場合は、表示されるドロップダウン・メニューから「ID」および「内容」のフィールドを選択します。

トランザクション形式

トランザクション形式データには、各トランザクションまたは項目に対応する個別のレコードがあります。例えば、顧客が複数の買い物をする時、それぞれの買い物が、顧客 ID によって関連付けられた関連項目を含む別個のレコードになります。これは、*ペーパー・ロール*形式と呼ばれることもあります。

顧客	購入物
1	ジャム
2	牛乳
3	ジャム
3	パン
4	ジャム
4	パン
4	牛乳

テーブル形式のデータ

テーブル形式のデータ（バスケットまたは真理値テーブル・データとも呼ばれる）では項目が別個のフラグで表され、各フラグのフィールドは特定の項目の有無を示します。各レコードは、関連付けられた項目の完全なセットを表します。フラグ型フィールドは、カテゴリーや数値フィールドにすることができますが、一部のモデルには、さらに具体的な要件がある場合があります。

顧客	ジャム	パン	牛乳
1	T	F	F
2	F	F	T
3	T	T	F
4	T	T	T

オプションのモデル設定

モデルの作成時に、以下に示すオプション設定を選択できます。選択したモデルのタイプおよびターゲットによっては、これらすべての設定が表示されない場合があります。既存モデルについてこれらの設定のいずれかを変更する場合は、変更を適用するために、モデルを再作成する必要があります。

自動的にデータをクリーンアップして準備し、モデル作成の信頼性を確保する。モデリングの処理速度、予測可能性、および信頼性を上げるために、データに関する問題を識別して修正します。インテリジェントなスクリーニングおよびサンプリング手法を通じて、欠落値や極端な値を処理し、必要に応じて新規属性を派生させ、パフォーマンスを向上させることなどにより、問題のあるフィールドや利用価値がなさそうなフィールドを選別します。新しいデータ・ソースを使用して最初にモデルを作成するときは、問題と修正を識別するために分析が行われるため、最初の実行時に限り、パフォーマンスが低下することがあります。以降の実行では、修正の適用は行われますが、データ・ソースに変更がなければ、分析が再度行われることはありません。この設定は、高度なユーザーが指定したカスタムのデータ準備設定を備えたモデルなど、一部のモデルについては無効にすることができます。

自動的にデータを区分して、検証およびテストのためのデータ・ソース作成時にモデル評価を実行できるようにする。このオプションを選択すると、モデルのトレーニング用とテスト用に、データが別々のサブセットまたはサンプルに分割されます。モデルの作成に使用したサブセットとは別のサブセットでテストを行うことで、その動作を他のデータ・セットに一般化する方法を推測できます。無作為に生成されるトレーニング・データ区分のサイズとテスト・データ区分のサイズにパーセント値を指定することもできます。

トレーニング・データ区分の最小サイズは 1、最大サイズは 100 です。テスト・データ区分の最小サイズは 0、最大サイズは 100 です。検証データ区分の最小サイズは 0、最大サイズは 100 です。モデルを正常に作成するには、パーセンテージの合計を 100 以下にする必要があります。

これらの区分化オプションは、アソシエーション・モデルを除くすべてのモデル・タイプに使用できます。

モデル作成用のモデル手法を選択。必要に応じて、「**選択**」をクリックして、モデルの作成時に含めるモデリング手法を選択します。選択したターゲットに基づき、使用可能なすべてのモデリング手法がリストされます。使用可能なモデリング手法（およびデフォルトで選択されるモデリング手法）は、選択したターゲットによって異なります。例えば、「**ディシジョン・リスト**」アルゴリズムが表示されるのは、フラグ・ターゲットの場合だけです。この機能は、予測モデルにのみ使用可能です。

モデリング手法について詳しくは、「*IBM SPSS Modeler アルゴリズム・ガイド*」、および IBM SPSS Modeler 製品に同梱の Web で入手可能なその他の資料を参照してください。

結合するモデルの最大数。保持および結合するモデルの最大数を設定できます。このオプションを 1 に設定すると、単一のモデルが作成および保持されます。ここに入力された値よりも作成されるモデルの数が少ない場合は、作成されるすべてのモデルが保持および結合されます。このオプションは、予測モデルにのみ使用可能です。

利益の基準 (2 値ターゲットを持つモデルのランク付けに使用)。モデルの作成前に、利益の基準を設定できます。ここで設定した値が、モデルの評価のデフォルトとして設定されます。このオプションは、フラグ・ターゲットが選択された予測モデルにのみ使用可能です。

使用する入力の指定。使用するフィールドを選択できます。これらは、通常、予測しようとしているものと何らかの実際的な関係があるもの（年齢、収入など）です。データ・セットの規模が大きい場合は、フィールドの数を制限することがモデルを単純化するための方法の 1 つです。顧客 ID や連絡先番号などのフィールドがデータに含まれている場合、これらは通常、モデル作成には役立たないため、選択しないようにする必要があります。他のデータと重複するフィールドも除外できます。

リンクが設定された入力フィールド (式) をクリックすると、その式の式ビューアーが開きます。式を編集するには、「データ」タブを表示します。詳しくは、43 ページの『式エディター』のトピックを参照してください。

使用する選択の指定。モデリング時に含める、または除外するレコードを指定します。既存のルールを検索することも、必要に応じて新規ルールを作成することもできます。詳しくは、47 ページの『選択ルールの定義』のトピックを参照してください。

さらに、グローバル選択が定義されている場合は、それらの選択がここに表示され、モデリング時にもこれらを適用するかどうかを指定できます。すべてのグローバル選択ルールを使用するか、グローバル選択ルールを使用しないかを選択できます。グローバル選択ルールの一部を選択することはできません。

セグメント化のオプション

予測ルール・モデルの作成時には、いくつかのセグメント化のオプションを使用できます。これらは、他のタイプのモデルで使用可能なオプションに加えて使用できます。

セグメントの最小サイズ。これらの設定は、セグメントの最小サイズを示し、2 つの値のうち大きい方が優先されます。例えば、パーセンテージの値が絶対値よりも大きい数値に相当する場合は、パーセンテージの設定が優先されます。

- **前のセグメントに対するパーセンテージ。**グループの最小サイズを、レコードのパーセンテージの形式で指定します。指定できる最小設定は 0、指定できる最大設定は 99.9 です。
- **絶対値。**グループの最小サイズを、レコードの絶対数で指定します。指定できる最小設定は 1 です。最大設定はありません。

最大属性数。セグメント・ルールあたりの条件の最大数を指定します。指定できる最小設定は 1 です。最大設定はありません。

属性の再使用を許可。有効にすると、前のサイクルで使用した属性も含め、各サイクルですべての属性を考慮できます。セグメントの条件はサイクル形式で作成され、各サイクルで 1 つの新規条件が追加されます。サイクルの数は、「**最大属性数**」設定を使用して定義します。

新しい条件の信頼区間。セグメント重要度のテストに対する確信度レベルを指定します。この設定は、セグメントあたりの条件数ルールとともに、返されるセグメントの数 (セグメントが返される場合) に大きく影響します。この値が大きいほど、返される結果セットは小さくなります。指定できる最小設定は 50、指定できる最大設定は 99.9 です。

自動クラスター・オプション

クラスターリング・モデルの作成時には、いくつかの自動クラスター・オプションを使用できます。これらは、他のタイプのモデルで使用可能なオプションに加えて使用できます。

評価フィールド。評価に使用するフィールドを指定します。評価フィールドを選択すると、IBM SPSS Modeler 自動クラスター作成ノードのランク付け方法が、「**評価フィールドの重要度**」に設定されます。評価フィールドを選択しない場合は、ランク付け方法がデフォルトのままになります。ここでは、評価フィールドとして分析フィールドを選択できます。

検索するクラスター数の範囲を目的に合わせて設定する。自動クラスター・モデルで検出するクラスター数に対し、任意の最小値と最大値の範囲を指定します。

アソシエーション・ルールとスコアリングのオプション

アソシエーション・モデルの作成時には、アソシエーション・ルールとスコアリングに関するいくつかのオプションを使用できます。これらは、他のタイプのモデルで使用可能な一部のオプションに加えて使用できます。

アソシエーション・ルール・オプション

- **範囲 (%)**。ルールの全体的な適用可能性を制御する、範囲の最小パーセンテージを指定します。例えば、値を 80% に設定すると、すべてのルールが少なくとも 80% のケースに適用されます。値が大きいほど、生成されるルールが少なくなります (ルールがまったく生成されない可能性もあります)。
- **確信度 (%)**。確信度の最小パーセンテージを指定します。
- **前提条件の最大数**。前提条件の最大数を指定します。
- **データは既に ID でソートされています (トランザクション形式専用)**。モデルの作成開始時にデフォルトで実行されるソートを無効にして ID の連続性を確保するには、このオプションを選択します。このオプションを選択すると、処理時間が短縮されます。データ形式がテーブル形式である場合、このオプションには効力がありません。

スコアリング・オプション

- **最大予測数**。含まれる予測の最大数を指定します。この設定は、「トップ」予測を生成するためにルール基準ドロップダウンとともに使用されます。このトップとは、最高レベルの確信度、サポート、リフトなどを示します。
- **ルール基準**。ルールの強度を判定するための基準を選択します。ルールは、アイテム・セットのトップ予測を返すために、ここで選択された基準の強度に従ってソートされます。使用可能な基準は以下のとおりです。
 - 「**確信度**」には、ルール条件支持度に対するルール支持度の比率が表示されます。これは、指定されたルール条件を持つレコードのうち、ルール予測も真であるものの比率を示します。例えば、トレーニング・データの 50% がパンを含んでいる (ルール条件支持度を指す) が、パンとチーズの両方を含むものが 20% (ルール支持度を指す) しかない場合には、ルール「パン -> チーズ」の予測は、ルール支持度/ルール条件支持度で、この場合は 40% になります。
 - 「**範囲 (%)**」には、ルール条件の支持度、すなわち、トレーニング・データに基づく、ルール条件が真である ID の比率が表示されます。例えば、トレーニング・データの 50% にパンの購入が含まれる場合は、ルール「パン -> チーズ」のルール条件支持度は 50% になります。
 - 「**ルール支持度 (%)**」には、ルール全体、ルール条件、およびルール予測が真であるレコードの比率が表示されます。例えば、トレーニング・データの 20% にパンとチーズの両方が含まれる場合、ルール「パン -> チーズ」のルール支持度は 20% になります。
 - 「**リフト**」には、ルール予測を持つ事前確率に対するルールの確信度の比が表示されます。例えば、母集団全体の 10% がパンを購入する場合、人々がパンを購入するかどうかを 20% の確信度で予測するルールのリフトは $20/10 = 2$ となります。別のルールで 11% の確信度で人々がパンを購入すると示された場合には、ルールのリフトは 1 に近くなります。つまり、そのルール条件を使用しても、ルール予測が得られる確率に大きな影響はない、ということです。一般に、リフトが 1 でないルールは、リフトが 1 に近いルールより効果的です。
 - 「**展開性 (%)**」は、ルール条件を満たすがルール予測を満たさないトレーニング・データの割合の尺度です。製品購入の分野では、基本的に、その条件を既に所有 (または購入) しているが、まだルール予測を購入していない顧客が、総顧客ベースの何パーセントにあたるかを示します。展開性の統計は「 $((\text{レコード数で表したルール条件支持度} - \text{レコード数で表したルール支持度}) / \text{レコード数}) * 100$ 」と定義されます。ここで、ルール条件支持度はルール条件が真であるレコードの数であり、ルール支持度はルール条件とルール予測のいずれも真であるレコードの数です。

- **トランザクション全体を検査。**以下のオプションから選択します。
 - 「**予測を表示しない**」を指定すると、ルール予測も提示されなくなります。スコアリングの目的が家具製品の提案である場合を例に挙げます。ここでは、既にショッピング・カートにダイニング・テーブルが入っている状態でもう 1 つ購入する可能性は低いと考えられます。そのような場合は、このオプションを選択してください。逆に、製品が生鮮食料品や使い捨て製品の場合 (チーズ、粉ミルク、ティッシュ・ペーパーなど) は、結果をあらかじめショッピング・カートに表示するルールが有用となる場合があります。後者のケースでは、「**予測を検査しない**」が最も有用となる場合があります。
 - 「**予測を表示する**」を指定すると、ルール予測もショッピング・カートに表示されるようになります。この方法は、既存の顧客やトランザクションに関する知見を得ようとしている場合に有用です。例えば、リフトが最高となるルールを突き止め、更に、どのような顧客がそれらのルールに合致するかを調べることができます。
 - 「**予測を検査しない**」の場合は、ショッピング・カートにルール予測を表示するかしないかにかかわらず、スコアリング時にすべてのルールが含まれます。
- **予測の繰り返しを許可。**スコアリング時に同じルール予測を持つ複数のルールを含めるには、このオプションを選択します。例えば、このオプションを選択すると、以下のルールのスコアリングが許可されます。

```
bread & cheese -> wine
cheese & fruit -> wine
```

スコアリング時に予測を繰り返さないようにするには、このオプションをオフにします。

- **一致しない項目を無視。**アイテム・セット内の追加アイテムの存在を無視するには、このオプションを選択します。例えば、ショッピング・カートの内容が [テント & 寝袋 & やかん] である場合にこのオプションが選択されていると、ショッピング・カートに余分なアイテム (やかん) があるにもかかわらず、ルール「テント & 寝袋 -> ガス・ストーブ」が適用されます。

モデルの評価

モデルの作成後、または完了したモデルを開いた後に、「**評価**」をクリックすると、分布、誤差の分布、一致行列、ゲイン、リフト、反応、利益、ROI などのグラフを取得して、実現可能な最大利益をシミュレートできます。使用可能なグラフのタイプとその他のオプションは、選択するモデル・タイプによって異なります。「**評価**」ダイアログを使用できるのは、選択したモデルが完了している場合のみです。アソシエーション・モデル・タイプでは「**評価**」は使用できません。

1. 上部のパネルで、評価データ・セットを選択します。モデルの作成時と同様、これは実際の反応がわかっている分析データ・セットでなければなりません (そうでなければ、実際の反応をモデルによって予測された反応と比べることができません)。
2. 評価に含めるモデルを変更するには、「**モデル選択の変更**」をクリックします。選択後に、以下に示す「**現在のモデル選択を、個々のモデル結果テーブルにコピーバック**」ボタンを、必要に応じてクリックすることもできます。これらのオプションは、予測モデル・タイプにのみ使用可能です。



図 6. 現在のモデル選択を、個々のモデル結果テーブルにコピーバック

3. 作成するグラフを選択します。
4. 利益をシミュレートする場合は、「**最大利益のシミュレート**」ボックスを選択し、利益計算の基盤となる母集団、予算、および最小値を入力します。詳しくは、65 ページの『**最大利益のシミュレート**』のトピックを参照してください。

5. 利益グラフまたは ROI グラフを要求する場合、または利益をシミュレートする場合は、コストおよび収益の値を指定します。これらの値は、各オファーのコストと、顧客がオファーに応じるたびに実現することが期待される収益の金額を、それぞれ示します。
6. 「ヒット」と見なす反応値を選択します。これは、モデルの作成に使用したターゲット・フィールドの値のうちのいずれかでなければなりません。例えば、良好な反応を示す場合は反応 = *true* です。
7. パフォーマンス・グラフを要求する場合は、「パフォーマンス・グラフについて個々のモデル評価を含める」オプションを選択することもできます。
8. 必要な場合は、「表形式の結果を含める (Include tabular results)」を選択します。これを選択した場合は、評価結果が返されるまでに要する時間が増加します。
9. 「実行」をクリックして結果を表示します。

分布図 (すべてのターゲット)

分布図とは、観測値と予測値を作図したもので、モデルの完了時にデフォルトで表示されます。「評価」ダイアログ・ボックスでは、テスト・データ・セット以外のデータ・セットの分布図を取得できます。範囲ターゲットの場合は、分布図が唯一の使用可能なグラフです。

選択できる追加のグラフは以下のとおりです。

誤差の分布図

誤差の分布図には、モデル結果に含まれるエラーのヒストグラムが表示されます。ここでエラーとは、予測値と実際の値との間の差異です。予測値が実際の値よりも大きい場合は正のエラーが発生し、予測値が実際の値よりも小さい場合は負のエラーが発生します。このグラフは、予測モデル・タイプの連続ターゲットにのみ使用できます。

一致行列グラフ

一致行列は、予測された結果と実際の結果のカウントを作図したものです。数値ターゲットの場合は、範囲が指定したピンの数に分割されます。

ゲイン・グラフ

ゲインは、モデルを使用することで得られるゲイン、つまり「リフト」を示すために使用され、有効なヒットの総数を基準とした各増分内のヒットの比率として定義されます。モデルを使用しない場合は、サンプル全体の期待される回答が、まっすぐな対角線によって作図されます。この場合、ある人が反応する可能性は他の人とまったく同じであるため、反応率は一定です。収益を倍にするには、倍の数の人にオファーを行う必要があります。これが曲線になっている場合は、反応する可能性が最も高い人を対象とすることによって、この結果をどの程度改善できるかが示されます。曲線が急であるほど、ゲインは大きくなります。

リフト・グラフ

リフト・グラフは、各増分に含まれるヒットと見なされるレコードのパーセンテージを、トレーニング・データ・セットに含まれるヒットの全体的なパーセンテージとともに作図したものです。

回答グラフ

回答グラフは、増分に含まれるヒットと見なされるレコードのパーセンテージを作図したものです。

利益グラフ

利益は、各レコードの収益から、レコードのコストを差し引いたものに相当します。分位の利益は、その分位内のすべてのレコードの利益を単純に合計したものです。収益はヒットにのみ適用されることが想定されていますが、コストはすべてのレコードに適用されます。詳しくは、『最大利益のシミュレート』のトピックを参照してください。

ROI グラフ

ROI (投資収益率) は、収益およびコストの定義に関連するという点で、利益と似ています。ROI では、各期間の利益とコストを比較します。

最大利益のシミュレート

モデルで予測された回答率に基づいて、特定の母集団について得られる最大利益をシミュレートすることができます。そのためには、コストと収益の値、および母集団のサイズを指定する必要があります。

最大利益を予測するには、以下の手順を実行します。

1. 予測のベースとなる母集団、つまりレコードの総数を指定します。注: この値は、モデルをスコアリングするときに、データ・ソースまたはサブセットの総レコード数に自動的に設定されます。
2. 利益シミュレーションに予算パラメーターを組み込む場合は予算の値を指定し、最小値を指定します。
3. 各レコードに関連付けられるコストを指定します。例えば、キャンペーン用販促資料の 1 人当たりの印刷費や送料などです。
4. 肯定的な反応から得られる予想収益を指定します。例えば、キャンペーンでのアイテムの小売原価などです。
5. 関心のある反応を選択します。
6. 「実行」をクリックします。

予測結果のグラフに、最大利益を得るための最良の結果が示され、モデルで特定された反応傾向に基づいて、この結果を得るためにターゲットとする必要があるレコードのパーセンテージが明示されます。

グラフの下にあるスライダー・バーを動かして、ターゲットとするレコードのパーセンテージを 0 から 100% の範囲で変更することで、利益に与える影響を確認できます。

許容できる数値を特定したら、レコードのスコアリングにその設定を使用するように指定できます。これを行うには、「スコアリングに使用」をクリックします。「スコア」タブの「傾向オプション」領域にある「上位 %」フィールドに、ターゲットとするレコードのパーセンテージが表示されます。

利益の比較

選択したレコードが適切な選択であることを検証するために、選択したレコード、無作為にグループ化した同じパーセンテージのレコード、全レコードのそれぞれから得られる見込み利益を比較できます。

表示される詳細には、ターゲットとしたレコードの数、見込まれる回答率、コスト、収入、および総利益などがあります。この情報を HTML 形式でエクスポートして、他のアプリケーションで使用することができます。

利益グラフ

選択したレコードから、予測される累積利益または非累積利益をグラフで表示できます。

注: このグラフは、評価を初めて実行するときに生成され、利益シミュレーション・セクションでスライダーを動かして変更を行っても自動的に更新されません。

モデル のテスト

「テスト」ダイアログ・ボックスを使用して、1 つ以上のサンプル・レコードがどのようにスコアリングされるかを確認することができます。結果は、テストが実行されるコンテキストによって異なり、モデルまたはルールによって返されるスコアが反映される場合があります。以下は、使用しているアプリケーション・タイプに応じて表示される結果の種類の例です。

- 少数の特定の顧客について、割り当てられたオファーを調べます。
 - 少数の顧客レコードについて、ダイレクト・メール キャンペーンに応答する確率の高い顧客を調べることができます。
 - 少数の特定の保険金請求について、割り当てられたアクションを調べます。
 - 少数の特定のケースについて、各ルールの評価を調べることができます。
1. 「モデリング」タブで、「テスト」をクリックします。「テスト」ダイアログが表示されます。
 2. 使用するデータ・ソースを選択するか、「レコード」セクションで「カスタム・データの指定」を選択して手動でテスト・レコードを入力します。
 3. 使用する選択があれば指定します。例えば、特定のレコードをテストに包含したり除外したりできます。既存のルールを検索することも、必要に応じて新規ルールを作成することもできます。詳しくは、47 ページの『選択ルールの定義』のトピックを参照してください。
 4. データ・ソースを使用する場合は、返されるレコードの数を指定します。上位のデータまたは無作為のサンプル・レコードから読み取られた最初の N レコード (最大 100 個) を返すことができます。トランザクション形式データを使用してアソシエーション・モデルをテストする場合は、入力データがレコード・レベルではなくトランザクション・レベルでサンプリングされます。
 5. カスタム・データを使用する場合は、「カスタム・データの指定」を選択し、「カスタム・レコードの追加」をクリックします。プロジェクト・データ・モデルで定義されているすべてのフィールドがリストされ、各フィールドに値を指定できるようになります。作業が終了したら「保存」をクリックします。
 6. 「表示するフィールドの選択」の下で、テスト出力に表示するフィールドを選択します。
 7. 「実行」をクリックします。テストの要約に、各レコードの成功した割り当てのリストが表示されます。
 8. 各行の先頭にある情報アイコンをクリックすると、その行の詳細が表示されます。

アプリケーションでのモデル・スコアの使用

モデルのスコアリングとは、決定に対する入力として使用できる結果または予測を得るために、データまたは母集団にモデルを適用することです。例えば、データベース内のすべての顧客をスコアリングして、ダイレクト・メール・キャンペーンの候補者を識別することや、コール・センターに連絡してくる顧客を 1 人ずつスコアリングして、その顧客に勧める製品またはサービスを判断することができます。スコアリング結果は、アプリケーションに応じて、データベース表やフラット・ファイルに書き込むことや、アプリケーションの決定を制御するセグメント・ルール、選択ルール、および割り当てルールの入力として使用することができます。

モデル・スコアは、以下の方法で使用できます。

- ルールに対する入力として。詳しくは、45 ページの『ルールの作成』のトピックを参照してください。

- バッチ・スコアリングをサポートするアプリケーションの場合は、モデル・スコアをデータベース表またはフラット・ファイルに書き込むことができます。詳しくは、69 ページの『データベース・テーブル、ファイル、Analytic Server、または、IBM Cognos BI サーバーのモデルのスコアリング』のトピックを参照してください。

モデルの出力フィールド

規定により、スコアリング出力フィールドには、ターゲット・フィールドの名前にカスタム接頭辞が付加された名前が付けられます。例えば、ターゲット・フィールドの名前が *Response* である場合は、出力フィールドの名前が *\$XF-Response* や *\$XFRP-Response* などとなります。以下で詳述するように、具体的なフィールドは、モデルのタイプ、および選択されたターゲット・フィールドの測定の尺度によって異なります。詳しくは、32 ページの『測定の尺度』のトピックを参照してください。

フラグ型対象。フラグ型対象を使用する自動化モデルは、以下のフィールドを出力します。

- **\$XF-<target>**。各レコードの予測値。フィールドに定義された「真」および「偽」の値で報告されます。
- **\$XFRP-<target>**。傾向スコア。各レコードの「真」値の尤度を示します。傾向スコアは確信度スコアとは異なり、通常はそのレコードの予測として報告されます。例えば、確信度の高い「偽」の予測は、反応しない尤度が高いことを示すため、低い傾向ということになります。

カテゴリー対象。カテゴリー対象 (名義型または順序型) を使用する自動化モデルは、以下のフィールドを出力します。

- **\$XS-<target>**。各レコードの予測値。
- **\$XSC-<target>**。予測に関連付けられた確信度値。

連続型対象。連続数値型対象を使用する自動化モデルは、以下のフィールドを出力します。

- **\$XR-<target>**。各レコードの予測値。
- **\$XRE-<target>**。予測の標準誤差。

予測ルール (対話型) モデル。予測ルール・モデル (フラグ型対象またはカテゴリー対象のみ) は、以下のフィールドを出力します。

- **\$D-<target>**。各レコードの予測結果。このフィールドに定義された「真」の値は望ましい反応を示し、Null はその他の値を示します。
- **\$DP-<target>**。各レコードの望ましい結果の確率。モデル作成時に結果テーブルのセグメントに表示される確率と一致します。詳しくは、55 ページの『予測ルール・モデルの作成』のトピックを参照してください。
- **\$DI-<target>**。レコードが属するセグメントを識別するインデックス番号 (モデル作成結果の左端の列に表示)。
- **\$DRP-<target>**。望ましい反応の尤度を示す傾向スコア。フラグ型対象の場合のみ得られます。

第 7 章 スコアリング・モデル

データベース・テーブル、ファイル、Analytic Server、または、IBM Cognos BI サーバーのモデルのスコアリング

モデルのスコアリングとは、モデルをデータ・サンプルまたは関心のある母集団に適用することです。例えば、既存の顧客を使用して、キャンペーンに反応する可能性が最も高い顧客を予測するモデルを生成し、このモデルを使用して、現在は顧客ではないが販売促進の送付先を含めたい人のレコードをスコアリングすることができます。

モデルをスコアリングする基本的な手順は以下のとおりです。

1. スコアリングするデータ・ソースおよびそのデータのサブセットを選択します。
2. スコアリングされたデータの宛先を選択します。
3. 出力する情報とそのスコアリング方法を選択します。
4. スコアリングするレコードを指定します。「スコア」ドロップダウンから、スコアリングの対象として、「すべてのレコード」、「上位 N %」、「上位 N レコード」、「最小/最大傾向」（フラグ型 対象でのみ選択可能）、または「最小スコア」（連続型 対象でのみ選択可能）を選択します。
5. 必要に応じて、予想される結果をプレビューし、データ概要を実行します。
6. 「ただちにスコアリング」をクリックして、指定した設定でモデルに対してスコアリングを実行します。完了したら、「スコアリング要約レポート」セクションを確認します。詳しくは、75 ページの『スコアリング要約レポート』のトピックを参照してください。

スコアは、データに追加された 1 つ以上のフィールドに書き込まれます。規定により、これらのフィールドには、ターゲット・フィールドに接頭辞が付加された名前が付けられます。例えば、ターゲット・フィールドの名前が *Response* である場合は、出力フィールドの名前が *\$XF-Response* や *\$XFRP-Response* となります。詳しくは、74 ページの『出力フィールドの選択』のトピックを参照してください。

注: ソース・データがデータベースから提供される場合、実行は、可能なときはいつでも（「SQL プッシュ・バック」と呼ばれるプロセスで）データベースにプッシュ・バックされます。通常、IBM Analytical Decision Management は、IBM SPSS Modeler と同じレベルの SQL プッシュ・バックをサポートしますが、以下の制限があります。

- モデルまたは外部ルールへの参照を持つ追加のフィールドまたは選択ルールがプロジェクトに含まれる場合、これらの参照はデータベースにプッシュ・バックされません。これにより、プロジェクト・ストリーム全体で達成可能なプッシュ・バックの程度が抑制されます。
- IBM SPSS Modeler Advantage 以外のアプリケーションの場合、プッシュ・バックは、選択ルールがアプリケーションで使用される限りにおいてのみ（「グローバル選択」タブ、および「定義」タブの最初のダイメンションのメンバー選択でのみ）有効になります。プロジェクト・データ・モデルのエンティティ ID として整数フィールドを選択した場合、プッシュ・バックの程度は増加することがあります。

データおよびサブセットの選択

「スコアリングするデータ・ソース」リストから、必要なデータ・ソースを選択します。

データ・サブセット

データ・ソース全体をスコアリングしたくない場合は、既存のルールを使用するか、またはこのモデル専用の新規ルールを作成して、データのサブセットを選択することができます。詳しくは、47 ページの『選択ルールの定義』のトピックを参照してください。

スコアリングするレコードの選択

傾向スコアを返すモデルでは、傾向スコアの値を使用して、スコアリングするレコードを選択できます。例えば、反応する傾向が最も高い 30% の顧客をスコアリングし、それらの顧客でメーリング・リストを作成することができます。傾向スコアは、対象に関する「真」値の尤度を示すもので、フラグ型対象を使用する予測モデル、および予測ルール (対話型) モデルで使用できます。(予測ルール・モデルの場合は、選択した反応が「真」値、それ以外がすべて「偽」となります。他のタイプのモデルでは、これらのオプションを使用できません)。

傾向に基づいてレコードを選択する場合は、以下の 3 つの方法を選択できます。

- **上位のパーセンテージ。** デフォルトでは、この方法で上位 10 パーセントが選択されます。このオプションでは、利益に基づいてスコアリング結果をシミュレートすることができます。詳しくは、65 ページの『最大利益のシミュレート』のトピックを参照してください。
- **上位のレコード数。** デフォルトでは、この方法で上位 1000 レコードが選択されます (存在する場合)。
- **傾向の最大値と最小値を指定。** 例えば、反応する傾向が 75% から 100% であるすべてのレコードを選択できます。

スコアリングの宛先の選択

スコアリング・データを、データベース、ファイル、Analytic Server、または、Cognos BI サーバーに出力するように指定することができます。サポートされるファイル・タイプには、テキスト (*.txt、*.csv)、Excel、IBM SPSS Statistics で使用される *.sav 形式、SAS、および IBM SPSS Data Collection でサポートされるファイルがあります。詳しくは、『宛先ファイル』のトピックを参照してください。

宛先データベース

スコアリング出力の送信先としてデータベースを選択すると、詳細を入力するように求められます。

1. 「**データベース名**」を入力するか、必要なデータベースを参照します。
2. スコアの格納方法を指定します。新しい表を作成することもできます。その場合は、表の名前を入力するように求められます。あるいは、既存の表にスコアを追加するか、既存のデータ・ソースを上書きすることができます。その場合は、関連する表またはデータ・ソースを選択するように求められます。
3. 必要に応じて、デフォルトの文字列サイズを変更します。

既存のデータベースへの追加または上書きを行う場合は、スコアリング・フィールドを既存の宛先内のフィールドにマップすることができます。詳しくは、73 ページの『スコアリング・フィールドのマップ』のトピックを参照してください。

宛先ファイル

スコアは、複数の異なるファイル・タイプに出力することができます。タイプごとに、既存のファイルを参照するか、スコアリング・データを追加する新しいファイルを作成します。

既存のファイルへの追加または上書きを行う場合は、スコアリング・フィールドを既存の宛先内のフィールドにマップすることができます。詳しくは、73 ページの『スコアリング・フィールドのマップ』のトピックを参照してください。

Excel ファイル

タイプ。 エクスポートする Excel ファイル・タイプを選択します。

フィールド名を含める。 フィールド名をワークシートの最初の行に含めるかどうかを指定します。

フラット・ファイル

スコアリング出力をフラット・ファイルに送信することを選択した場合は、詳細を入力してファイルの内容を制御できます。

書き込みモード。 「上書き」を選択すると、指定されたファイル内のすべての既存データが上書きされます。「追加」を選択すると、既存ファイルの最後に出力が追加され、ファイルに含まれているすべてのデータが保持されます。

- **フィールド名を含める。** このオプションを選択すると、フィールド名が出力ファイルの最初の行に書き込まれます。このオプションは、「上書き」書き込みモードの場合にのみ使用できます。

各レコードの後に改行。 このオプションを選択すると、出力ファイルに、各レコードが改行されて書き込まれます。

フィールド区切り文字。 生成されたテキスト・ファイルでフィールド値の間に挿入する文字を指定します。「コンマ」、「タブ」、「スペース」、および「その他」をオプションとして使用できます。「その他」を選択する場合は、使用する区切り文字をテキスト・ボックスに入力してください。

シンボル値の引用符。 記号フィールドの値に使用する引用符のタイプを指定します。「なし」（値は引用符で囲まれません）、「単一引用符 (')」、「二重引用符 (")」、および「その他」をオプションとして使用できます。「その他」を選択する場合は、使用する引用文字をテキスト・ボックスに入力してください。

「エンコード」。 使用されるテキスト・エンコード方式を指定します。システム・デフォルトまたは UTF-8 を選択できます。

小数点記号。 データ内で小数点がどのように表示されるかを指定します。

- **ピリオド (.)。** ピリオド文字が、小数点として使用されます。
- **コンマ (,)。** コンマ文字が、小数点として使用されます。

IBM SPSS Statistics データ・ファイル

フィールド名のエクスポート。 IBM SPSS Statistics の .sav ファイルにエクスポートする際に変数名とラベルを処理する方法を指定します。無効な名前は、自動的に無効な文字を置き換えて修正されるので注意してください。

- **名前および変数ラベル。** フィールド名とフィールド・ラベルの両方をエクスポートする場合に選択します。名前は IBM SPSS Statistics 変数名としてエクスポートされ、ラベルは IBM SPSS Statistics 変数ラベルとしてエクスポートされます。
- **変数ラベルとしての名前。** フィールド名を IBM SPSS Statistics で変数ラベルとして使用する場合に選択します。無効な IBM SPSS Statistics 名が作成されないようにする場合は、代わりに「名前および変数ラベル」を選択してください。

SAS データ・ファイル

タイプ。作成する SAS ファイル形式を選択します。3 つの SAS ファイル形式、つまり「SAS for Windows/OS2 (*.sd2)」、「SAS for UNIX (*.ssd)」、「SAS バージョン 7/8/9 (*.sas7bdat)」から選択できます。

フィールド名のエクスポート。SAS で使用するフィールド名とラベルをエクスポートするためのオプションを選択します。

- **名前および変数ラベル。**フィールド名とフィールド・ラベルの両方をエクスポートする場合に選択します。名前は SAS 変数名としてエクスポートされ、ラベルは SAS 変数ラベルとしてエクスポートされます。
- **変数ラベルとしての名前。**フィールド名を SAS で変数ラベルとして使用する場合に選択します。無効な SAS 名が作成されないようにする場合は、代わりに「名前および変数ラベル」を選択してください。

Analytic Server 宛先

スコアリング出力の送信先として Analytic Server を選択すると、詳細を入力するように求められます。

1. 「**データベース・ソース**」を指定します。
2. スコアの格納方法を指定します。あるいは、既存の表にスコアを追加するか、既存のデータ・ソースを上書きすることができます。その場合は、関連する表またはデータ・ソースを選択するように求められます。
3. スコアリング出力に含める出力フィールドを選択します。
4. スコアリングするレコードを指定します。「スコア」ドロップダウンから、スコアリングの対象として、「**すべてのレコード**」、「**上位 N %**」、「**上位 N レコード**」、「**最小/最大傾向**」（フラグ型 対象でのみ選択可能）、または「**最小スコア**」（連続型 対象でのみ選択可能）を選択します。

Cognos BI サーバー宛先

スコアリング出力の送信先として Cognos BI サーバーを選択すると、詳細を入力するように求められます。

注: エクスポートできるのはリレーショナル・データのみで、OLAP データはエクスポートできません。

Cognos BI にデータをエクスポートするには、以下を指定する必要があります。

- **Cognos 接続。**Cognos BI サーバーへの接続。
- **ODBC 接続。**Cognos BI サーバーが使用する Cognos データ・サーバーへの接続。

上記の接続は、同じデータベースを指している必要があります。また、ODBC に接続する Cognos サーバーのユーザー名とパスワードは、ODBC の詳細と同じでなければなりません。

データ・サーバーには実際のデータをエクスポートし、Cognos BI サーバーにはパッケージ・メタデータをエクスポートします。

1. Cognos BI サーバーの「**宛先**」を選択し、「**接続**」をクリックします。サーバー接続の詳細を入力するように求められます。
 - a. データのインポート元またはエクスポート先となる IBM Cognos サーバーの「**サーバーの URL**」を入力します。使用する URL が不明な場合は、Cognos システムの管理者に問い合わせてください。
 - b. 接続する際の「**モード**」を選択します。特定のユーザー（管理者など）としてログインするには、「**資格情報の設定**」を選択し、IBM Cognos の名前空間、ユーザー名、およびパスワードを入力します。

- サーバーへのログオンに使用する、IBM Cognos セキュリティー認証プロバイダーの「**名前空間**」を入力します。認証プロバイダーは、ユーザー、グループ、および役割の定義と保守に使用されるほか、認証プロセスの制御に使用されます。
 - サーバーへのログオンに使用する IBM Cognos の「**ユーザー名**」を入力します。
 - 指定したユーザー名に関連付けられている「**パスワード**」を入力します。
- c. ユーザー資格情報なしでログインする場合は、名前空間、ユーザー名、およびパスワードの各フィールドを指定できないため、「**匿名接続の使用**」を選択します。

注: 一部のサーバー接続では、匿名接続を使用できません。

2. 「**データ・ソース**」を選択します。
3. エクスポート・パッケージの作成場所となる Cognos BI サーバー上の「**フォルダー**」のパスと名前を入力します。または、必要なフォルダーを参照することができます。
4. エクスポートされたメタデータを格納するパッケージの「**パッケージ名**」を入力します。これは新規のパッケージでなければなりません。既存のパッケージにはエクスポートできません。詳しくは、39 ページの『Cognos オブジェクト詳細の選択』のトピックを参照してください。
5. ODBC 宛先を選択する場合は、「**データベース名**」を入力するか、必要なデータベースを参照します。
6. スコアの格納方法を指定します。新しい表を作成することもできます。その場合は、表の名前を入力するように求められます。あるいは、既存の表にスコアを追加するか、既存のデータ・ソースを上書きすることができます。その場合は、関連する表またはデータ・ソースを選択するように求められます。

既存のデータベースへの追加または上書きを行う場合は、スコアリング・フィールドを既存の宛先内のフィールドにマップすることができます。詳しくは、『スコアリング・フィールドのマップ』のトピックを参照してください。

スコアリング・フィールドのマップ

既存のデータベース表またはファイルに対してスコアの追加または上書きを行う場合は、スコアリング・フィールドを既存の表内またはファイル内のフィールドにマップする必要があります。出力のデータ型と関連するマップ・フィールドのデータ型が同じでなければなりません。

- 既存のデータベース表にスコアを書き込む場合は、新しいスコアを上書きするか追加するかにかかわらず、スコアリングを進めるために、すべてのフィールドを既存のフィールドにマップする必要があります。必要であれば、「**フィールドのマップ**」をクリックして、残りのフィールドに出力を指定します。
 - 既存のファイルに書き込む場合は、出力ファイルのオプションで「**追加**」を選択した場合にのみマップが必要となります。既存のファイルを上書きする場合は、新規のファイルが既存のファイルに置き換わるだけなので、マップは不要です。
1. 「**フィールドのマップ**」をクリックします。ダイアログが開き、システムによって、フィールドのマップについて最適な推測が行われます。任意の時点で、「**リセット**」をクリックしてシステムの最適な推測に戻すことができます。
 2. 有効な出力の宛先を選択し、「**マップ**」をクリックして必要なスコアリング・フィールドにマップします。
 3. 作業が完了したら、「**保存**」をクリックします。

出力フィールドの選択

「スコア」タブで、「出力フィールドの指定」を選択してスコアリング出力に含めるフィールドを選択します。通常は、実際のスコアに加えて、各レコードを識別するためのフィールド（顧客 ID フィールドなど）が少なくとも 1 つは必要になります。オプションとして、モデル作成で使用された入力フィールドの全部または一部を含めることもできます。

モデルの出力フィールド

規定により、スコアリング出力フィールドには、ターゲット・フィールドの名前にカスタム接頭辞が付加された名前が付けられます。例えば、ターゲット・フィールドの名前が *Response* である場合は、出力フィールドの名前が *\$XF-Response* や *\$XFRP-Response* などとなります。以下で詳述するように、具体的なフィールドは、モデルのタイプ、および選択されたターゲット・フィールドの測定の尺度によって異なります。詳しくは、32 ページの『測定の尺度』のトピックを参照してください。

フラグ型対象。フラグ型対象を使用する自動化モデルは、以下のフィールドを出力します。

- **\$XF-<target>**。各レコードの予測値。フィールドに定義された「真」および「偽」の値で報告されます。
- **\$XFRP-<target>**。傾向スコア。各レコードの「真」値の尤度を示します。傾向スコアは確信度スコアとは異なり、通常はそのレコードの予測として報告されます。例えば、確信度の高い「偽」の予測は、反応しない 尤度が高いことを示すため、低い傾向ということになります。

カテゴリ対象。カテゴリ対象（名義型または順序型）を使用する自動化モデルは、以下のフィールドを出力します。

- **\$XS-<target>**。各レコードの予測値。
- **\$XSC-<target>**。予測に関連付けられた確信度値。

連続型対象。連続数値型対象を使用する自動化モデルは、以下のフィールドを出力します。

- **\$XR-<target>**。各レコードの予測値。
- **\$XRE-<target>**。予測の標準誤差。

予測ルール（対話型）モデル。予測ルール・モデル（フラグ型対象またはカテゴリ対象のみ）は、以下のフィールドを出力します。

- **\$D-<target>**。各レコードの予測結果。このフィールドに定義された「真」の値は望ましい反応を示し、Null はその他の値を示します。
- **\$DP-<target>**。各レコードの望ましい結果の確率。モデル作成時に結果テーブルのセグメントに表示される確率と一致します。詳しくは、55 ページの『予測ルール・モデルの作成』のトピックを参照してください。
- **\$DI-<target>**。レコードが属するセグメントを識別するインデックス番号（モデル作成結果の左端の列に表示）。
- **\$DRP-<target>**。望ましい反応の尤度を示す傾向スコア。フラグ型対象の場合のみ得られます。

モデルのスコアリング

スコアリングの詳細を指定したら、モデル・スコアのサンプルをプレビューすることができます。これにより、データ・セット全体がスコアリングされるのを待たなくても、スコアリングの結果が予想どおりであることを簡単に確認できるため、大きなデータ・セットがある場合は特に便利です。

注: データが無効または不完全と判断された場合は、プレビューを試みた時点で警告が表示されます。

プレビューでは、レコードを行、変数を列とする表形式でデータが表示されます。

スコアリング

スコアリングするレコード数を設定するには、「更新」アイコンをクリックします。これにより、データ・ソース内の総レコード数および指定したオプション設定を考慮に入れて、モデルの作成に使用されるレコードの数が表示されます。

「ただちにスコアリング」をクリックすると、タブの下部に進行状況表示バーが表示されます。これが表示されている間は、いつでもスコアリングを中止できます。

注:

データが無効または不完全と判断された場合は、スコアリングを試みた時点で警告が表示されます。

データベースにスコアリングする場合、実行は、可能なときはいつでもデータベースにプッシュ・バックされます (多くの場合、「SQL プッシュ・バック」と呼ばれます)。モデルへの参照を含むルールがプロジェクト内にある場合、実行はデータベースにプッシュ・バックされないことに注意してください。

データが正常にスコアリングされると、結果の格納場所 (指定したデータベースまたはファイル) がメッセージで通知され、「スコアリング要約レポート」セクションに要約が表示されます。詳しくは、『スコアリング要約レポート』のトピックを参照してください。スコアリング出力をフラット・ファイルに送信することを選択し、かつ管理者がその機能を有効にした場合は、必要に応じてそのファイルのローカル・コピーをダウンロードできます。

スコアリング要約レポート

プロジェクトをスコアリングすると、「スコアリング要約レポート」セクションに、データのスコアリング先、ファイル・サイズ (該当する場合)、データベースと表名 (該当する場合)、モデル名またはプロジェクト名、スコアリングされたレコードの数、および場合によってはモデルを要約したグラフとテーブル (またはそのいずれか) などの詳細が表示されます。「スコアリング要約レポート」セクションは、IBM SPSS Modeler Advantage で表示されるほか、「ただちにスコアリング」機能を備えたアプリケーションの「展開」タブに表示されます。

スコアリング出力をフラット・ファイルに送信することを選択した場合は、必要に応じてそのファイルのローカル・コピーをダウンロードすることもできます。管理者は、ブラウザー・ベースの IBM SPSS Collaboration and Deployment Services Deployment Manager を使用して、この機能を有効または無効にできるほか、許容される最大ファイル・サイズを増減することができます。「ファイルのダウンロード」ボタンがアクティブでない場合は、ファイル・サイズが管理者の設定した最大サイズを超えているか、または管理者がこの機能を無効にしています。この機能は、デフォルトで無効になっています。

第 8 章 アプリケーションの管理

管理者は、ビジネス・ユーザーが IBM Analytical Decision Management アプリケーションと対話する方法を制御できます。このレベルの制御を行うことで、ビジネス・ユーザーのジョブはより簡単になるため、アプリケーションをより素早く効率的に使用して、ビジネス問題の解決することができます。また、管理者が特定の機能に対するアクセスを制限することもできます。

必要に応じて、IBM SPSS Collaboration and Deployment Services Deployment Manager を使用して特定のユーザーまたはグループに管理権限を付与してください。「*Decision Management の管理*」アクションを付与されているユーザーのみが、このセクションで説明されている機能にアクセスできます。

1. IBM SPSS Collaboration and Deployment Services Deployment Manager を開きます。
2. 「ツール」メニューから、「**サーバー管理**」を選択します。
3. 「サーバー管理」タブで、サーバーに接続します。
4. 「**ユーザーとグループ**」で、「**ローカル・ユーザー・リポジトリ**」を選択します。
5. 「**新規グループ**」をクリックして、管理者用の新しいグループと、ビジネス・ユーザー用の新しいグループを作成します。以下に例を示します。
 - Decision Management 管理者
 - Decision Management ユーザー

また、必要に応じて、この時点で新規ユーザーを作成することもできます。グループを作成するときに、各グループに該当するユーザーを追加します。例えば、管理者グループのメンバーは、組織の IBM Analytical Decision Management アプリケーションを管理する権限を付与されます。

6. 「役割」に移動して「**新規役割**」をクリックし、新規管理役割と新規ビジネス・ユーザー役割を作成します。以下に例を示します。
 - IBM Analytical Decision Management 管理者
 - IBM Analytical Decision Management ユーザー

新規役割の作成時に、少なくとも以下のアクションを役割に割り当てます。

表 5. 役割のアクション：

役割	必要なアクション
Decision Management 管理者	コンテンツとフォルダーへのアクセス
	Decision Management の管理
	ジョブの編集
	レポートの動的実行
	モデルのスコアリング
	処理の送信
	すべてのバージョンの表示

表 5. 役割のアクション (続き):

役割	必要なアクション
Decision Management ユーザー	コンテンツとフォルダーへのアクセス
	ジョブの編集
	レポートの動的実行
	モデルのスコアリング
	処理の送信
	最新の表示または すべてのバージョンの表示*

*注: ユーザーが自分で作成しなかったプロジェクト・バージョンにアクセスする必要がある場合、最新の表示 または すべてのバージョンの表示 アクションのいずれかが必要です。アクションがなければ、現在のユーザーが最新バージョンへのアクセス権を持っていない場合、プロジェクトは読み取り専用モードで開きます。

- Decision Management 管理者役割を選択し、「ユーザーとグループの編集」をクリックします。作成した Decision Management 管理者グループを追加します。
- Decision Management ユーザー役割を選択し、「ユーザーとグループの編集」をクリックします。作成した Decision Management ユーザー・グループを追加します。

一般オプション

IBM Analytical Decision Management ユーザー・インターフェースのすべてのタブで、管理者は特定の一般オプションを使用できます。管理者はこれらのオプションを使用して、ビジネス・ユーザーに対してタブを非表示にしたり、ロックしたりすることができます。

- 各タブの上部に、「管理者以外のユーザーに「<tab name>」タブを表示しない」 (<tab name> はタブの名前) があります。これを使用すると、そのタブ全体がビジネス・ユーザーに対して非表示になります。例えば、ユーザーに対してアプリケーションの「データ」タブを非表示にする場合、そのアプリケーションにログオンして、「データ」タブに移動し、「管理者以外のユーザーに「データ」タブを表示しない」を選択します。
- 各タブの上部に「すべての <tab name> オプションのロック」 (<tab name> はタブの名前) があります。これを使用すると、タブ上のすべてのコントロールがロックされ、ビジネス・ユーザーに対して読み取り専用になります。例えば、ビジネス・ユーザーに対して、現在のデータ設定の表示は行えるが、変更は行えないようにする場合、アプリケーションの「データ」タブでこのオプションを設定できます。この設定は、ビジネス・ユーザーがジョブを簡素化できるように、前もって特定の事柄を設定する必要がある場合にも非常に便利です。
- 管理者は、使用可能な場合はパッドロック・アイコンをクリックして、各タブの個々の設定をロックしたり、ロック解除したりすることもできます。



例えば、「定義」タブがあるアプリケーションで、管理者はディメンション階層ツリーをロックして、ビジネス・ユーザーがそれを変更できないように設定できます。パッドロック・アイコンをクリックすると、それに関連付けられたコントロールがロックまたはロック解除されます。ビジネス・ユーザーには、管理者がロックしたコントロールにはロックされたパッドロック・アイコンが表示されますが、ロック解除されたパッドロック・アイコンは表示されません。

ロックされた個々の項目をグレーの背景やくすんだ色で表示して、選択できないことを示し、ユーザーがそれらを選択しようとする、エラー・メッセージが表示されるようにすることもできます。タブ全体がロックされている場合、そのことを示すメッセージがタブの上部に表示されます。

アプリケーション・ショートカットの制御

「アプリケーション起動ページ」には、ビジネス・ユーザーがアプリケーションを開いたり、アプリケーションの特定のラベル付きバージョンにカスタム・ショートカットを追加したりするための、簡単な方法が用意されています。管理者は、すべてのユーザーの起動ページに表示されるアプリケーション・ショートカットを指定できます。管理者権限でログオンした後、以下を実行します。

- 必要な場合は、アプリケーション・ショートカットのパッドロック・アイコンをクリックして、ビジネス・ユーザーがそれを起動ページから削除できないようにします。この場合でも、ビジネス・ユーザーは他のアプリケーションのショートカットを追加および削除できます。ユーザーには、ロックされたショートカットには、閉じるアイコンではなく、ロックされたパッドロック・アイコンが表示されます。詳しくは、21 ページの『アプリケーションの起動』のトピックを参照してください。

ホーム・ページの非表示設定

各アプリケーションには、ホーム・ページのグラフィックによって表される、段階的なワークフローが用意されています。グラフィックをクリックすると、ユーザー・インターフェースのそのセクションにジャンプします。ただし、ビジネス・ユーザーによっては、アプリケーション・ワークフロー内のすべてのステップにアクセスしなくてもよい場合もあります。

管理者は、ビジネス・ユーザーに対してホーム・ページを非表示にすることができます。例えば、特定のアプリケーションのビジネス・ユーザーは、「レポート」タブしか使用しない場合があります。このような場合、管理者はそのようなビジネス・ユーザーに対して「レポート」タブを除く他のすべてのタブと、ホーム・ページを非表示にすることができます。このようにすれば、ビジネス・ユーザーがログオンすると「レポート」タブが直接開かれ、その他のタブは表示されなくなります。

1. ホーム・ページで、「**アプリケーション・ホーム・ページの非表示 (Hide Application home page)**」を選択します。
2. ドロップダウンで、ユーザーが最初にアプリケーションを開いたときに表示されるタブを指定します。

選択したこのタブを後で非表示にすると、アプリケーション内で次に非表示に設定されていないタブが表示されます。

データ・オプションのロック

データ

「データ」タブでは、アプリケーションが分析、シミュレーションとテスト、スコアリング、またはその他のオペレーションに使用するデータ・セットを定義します。これらのデータ・ソースは、サーバーに相対的に定義されます。したがって、データ・ファイルや ODBC ソース等へのパスは、すべてサーバーに相対的になります。

管理者は、ビジネス・ユーザーがアプリケーションの使用を開始する前に、データの一部またはすべてをセットアップしておくこともできます。管理者は、以下を行うことができます。

- 新しいデータ・ソースを作成し、作成後にそれをロックする場合に、「**データ・ソースのロック**」を選択する。

- テーブル内または「データ・ソース・エディター」ダイアログ内でデータ・ソースの横にあるパッドロック・アイコンをクリックして、個々のデータ・ソースをロックまたはロック解除する。
- 「プロジェクト・データ・モデル」ドロップダウンでデータ・モデルを選択し、パッドロック・アイコンをクリックすることにより、アプリケーションで使用するデータ・モデルを設定して、ロックする。

ビジネス・ユーザーは「プロジェクト・データ・セット」セクションから「マイ・データ・セット」セクションに、ロックされたデータ・ソースをコピーできますが、そのデータ・ソースは両方のセットでロックされたままになります。詳しくは、29 ページの『第 4 章 データ・ソースの管理』のトピックを参照してください。

- プロジェクト・データ・モデルがデータ・ビューの場合、「階層データに基づいて ODM ルールの使用を許可 (Allow use of ODM rules based on hierarchical data)」ドロップダウンを使用して、階層データ・モデルを持つ IBM Operational Decision Management (ODM) ルールの使用法を制御します。このようなルールを使用するとパフォーマンスの低下を招く恐れがあります。次のオプションはドロップダウンから利用できます。
 - 「許可しない (Don't allow)」は、最高のパフォーマンスを実現する場合に推奨されます。プロジェクトでの複雑なデータ構造での ODM ルールの使用は禁止するようにしてください。アプリケーション内でユーザーが複雑なルールを選択すると、エラーが表示されます。
 - 「モデリング・タブ上のみ (On Modeling tab only)」では、「モデリング」タブでのモデルの作成時に複雑なルールの参照が許可されます。これにより、柔軟性は向上しますが、パフォーマンスの低下が発生します。このオプションは、IBM SPSS Modeler Advantage にのみ適用されます。

スコアリング宛先オプションのロック

「ただちにスコアリング」セクション（「展開」タブまたは「スコア」タブ）を使用するように構成されたアプリケーションの場合、ビジネス・ユーザーがスコア・データを（データベースまたはファイルに）保存するための宛先を指定できるセクションがあります。例えば、スコアをデータベースに保存する場合、データベース名とスコアの保存方法（新しい表をデータベースに作成するなど）を指定します。

管理者は、スコアリングの宛先オプションをロックすることができます。例えば、特定のデータベースに存在するテーブルにのみスコアを追加し、ビジネス・ユーザーには、既存データの上書き、既存の表への新しいデータの追加、新しい表の作成、またはファイルへのデータの保存などを行うためのオプションを付与しないようにできます。

- すべてのスコアリング宛先オプションをロックするには、「展開」タブの「宛先」ドロップダウンの横にあるパッドロック・アイコンをクリックします。ビジネス・ユーザーは宛先オプションを表示することはできますが、編集することはできません。

外部ルールの作成に使用するプロジェクト・メタデータのダウンロード

IBM Operational Decision Management などのビジネス・ルール管理システムで作成されたルールは、現在の IBM Analytical Decision Management プロジェクトで使用されているデータ・モデルをサポートするように開発されたことを条件として、IBM Analytical Decision Management アプリケーションで参照および使用することができます。また、このアプリケーションで使用できるように Web サービスとして展開できます。これを行うために、IBM Analytical Decision Management 管理者は、関連するメタデータをダウンロードする必要があります。メタデータをダウンロードするには 2 つの方法があります。データ・モデルの複雑さと、作成される可能性のあるルールの複雑さは、それに応じて異なります。

IBM Analytical Decision Management からのプロジェクト・メタデータのダウンロード

管理者は、IBM Analytical Decision Management で「メタデータのダウンロード」アイコンをクリックして、現在のプロジェクトのメタデータを含む .ZIP ファイルをダウンロードできます。ダウンロードしたファイルは、管理者がダウンロードを実行したシステムにローカルで保存することができます。



図7. 「メタデータのダウンロード」アイコン

この方法でダウンロードされたメタデータにより、プロジェクト・データ・モデルに基づくルールの作成がサポートされます。プロジェクト・データ・モデルは (階層型ではなく) フラットになります。そのため、このメタデータに基づいて作成されたすべてのルールは、フラット・データ・モデルのみを利用することができます。プロジェクト・データ・モデルのデータ・ソースのタイプにかかわらず、このフラット・データ・モデルに基づいて作成されたルールは、IBM Analytical Decision Management プロジェクトで参照することができます。

.ZIP ファイルには以下のファイルが含まれています。

- **XML スキーマ定義 (*.XSD)**。現在の IBM Analytical Decision Management プロジェクトのデータ・モデル内にあるフィールド・データ型の定義が含まれています。このファイルは、現在のプロジェクトで使用するルールを開発するために、IBM Operational Decision Management Rule Designer などの外部開発ツールにインポートすることができます。
- **Web サービス記述言語テンプレート (*.WSDL)**。データ・モデルを使用する Web サービスの開発を支援するために組み込まれています。.WSDL ファイルは、IBM Analytical Decision Management アプリケーションと外部ルール・サービスの間でメッセージを仲介する Web サービスを構築するために使用できるテンプレートとして提供されます。仲介を行う Web サービスに関する詳細は、インテグレーターが手動で入力できます。

IBM SPSS Collaboration and Deployment Services 分析データ・ビューのデータ・モデル・メタデータのダウンロード

IBM SPSS Collaboration and Deployment Services Deployment Manager では、特定の分析データ・ビューのメタデータをエクスポートすることができます。

この方法でダウンロードされたメタデータにより、分析データ・ビューの論理データ・モデルに基づくルールの作成がサポートされます。分析データ・ビューの論理データ・モデルは、階層型にすることができます。IBM Analytical Decision Management データ・ビュー・ソースでは、XML 表現による各レコードの階層データの生成がサポートされており、このデータは実行時に外部ルールに渡すことができます。このメタデータに基づいて作成されたルールは、複雑な場合があるため、階層データ・モデルが利用されます。

IBM Analytical Decision Management の場合、階層データ・モデルに基づく外部ルールは、プロジェクト・データ・モデルがデータ・ビュー・ソースを使用するプロジェクトでのみ参照することができます。互換性を維持するために、データ・ビュー・ソースおよび外部ルールは、同じ分析データ・ビューの論理データ・モデルに基づく必要があります。参照先の分析データ・ビュー・インスタンスに基づく階層データを生成できるのは、データ・ビュー・ソースに限られます。言い換えると、外部ルールを使用するには、IBM Analytical Decision Management プロジェクトのプロジェクト・データ・モデルが、ルール作成に使用されるプロジェクト・データ・モデルに一致する必要があります。これは、分析データ・ビューによって可能になります。

分析データ・ビューの論理データ・モデルのメタデータを含む *.ZIP* ファイルをダウンロードするには、以下の手順を実行します。

1. IBM SPSS Collaboration and Deployment Services Deployment Manager で、リポジトリから分析データ・ビューを開き、「論理」タブを選択します。
2. テーブルを右クリックして、XOM メタデータをエクスポートします。
3. ファイルを保存する場所を選択して、「保存」をクリックします。

.ZIP ファイルには以下のファイルが含まれています。

XML スキーマ定義 (*.XSD)。 2 つ以上の XML スキーマ定義が含まれます。 1 つのスキーマには、ルール回答タイプに役立つ可能性のある定義が含まれます。残りのスキーマには、テーブルの定義 (分析データ・ビューの論理データ・モデルの階層関係の詳細やフィールド・タイプなど) が含まれます。

プロジェクトのプロジェクト・データ・モデルで同じ分析データ・ビュー・インスタンスに基づくデータ・ビュー・ソースが使用されている場合、これらのスキーマ・ファイルを IBM Operational Decision Management Rule Designer などの外部開発ツールにインポートすると、そのようなプロジェクトで使用されるルールを開発することができます。

IBM Analytical Decision Management で使用する外部ルールの作成については、「アプリケーション・デザイナーズ・ガイド」を参照してください。

これらの複雑な外部ルールはパフォーマンスの低下を招くおそれがあるため、管理者がその使用を制限している場合があることに注意してください。詳細については、79 ページの『データ・オプションのロック』を参照してください。

XML スキーマ定義の編集

IBM Analytical Decision Management からダウンロードされた *.XSD* ファイルには、2 つの複合データ型が含まれています。1 つは、IBM Analytical Decision Management プロジェクト・データ・モデルを表し、IBM Operational Decision Management ルール・プロジェクトの入力パラメーターに使用されます。もう 1 つは、ルールで返されるデータ型を表し、IBM Operational Decision Management ルール・プロジェクトの出力パラメーターに使用されます。

IBM SPSS Collaboration and Deployment Services Deployment Manager の分析データ・ビューのメタデータには、*.XSD* ファイルが 2 つ以上含まれています。その内の一つは、分析データ・ビューの論理データ・モデルに関連するものであり、残りは、ルールで返されるデータ型を表す複合型を含むものです。前者は、IBM Operational Decision Management ルール・プロジェクトの入力パラメーターに使用する必要があり、後者は、IBM Operational Decision Management ルール・プロジェクトの出力パラメーターに使用する必要があります。

エキスパート・ユーザーは、生成された *.XSD* スキーマ定義を編集して、作成する特定のルールに合わせて調整することができます。場合によっては、ルールを作成する前に、関連する *.XSD* スキーマ定義で定義されている入力関連データ型と回答関連データ型をユーザーが編集することもあります。例えば、作成したルールで、Null 値を持つ入力フィールドを必要に応じて処理できるようにするには、エキスパート・ユーザーが入力パラメーター関連のスキーマを編集することが必要になる場合があります。これは、どの入力パラメーターのフィールド/要素が必須であり、またオプションであるのかを *.XSD* スキーマ定義を介して指定することにより行うことができます。

入力パラメーターのデータ型の編集

各フィールド関連要素の出現可能数は、XSD スキーマでは明示的に設定されません (つまり、属性 minOccurs と maxOccurs は設定されず、両属性のデフォルト値は 1 です)。このため、編集前のスキーマを使用して IBM Operational Decision Management のルールが生成された場合は、それらのルールで入力パラメーター・データ型のすべてのフィールドが必要となります。

プロジェクト・データ・モデル関連の入力パラメーター・タイプの編集

IBM Analytical Decision Management プロジェクトのメタデータをダウンロードすることにより、入力パラメーター・タイプを定義しているスキーマが生成された場合、入力パラメーターのデータ・モデルはフラットになります。この場合、IBM Analytical Decision Management は、実行時にルール・サービスで必要なペイロードを直接作成します。これにより、エキスパート・ユーザーは、より自由に入力パラメーター・タイプを編集し、なおかつ互換性を維持することができます (つまり、ルール作成に不要なフィールド/要素を削除したり、フィールド/要素がオプションであることを示したりできます)。

分析データ・ビューの論理データ・モデル関連の入力パラメーター・タイプの編集

分析データ・ビューの論理データ・モデルを IBM SPSS Collaboration and Deployment Services Deployment Manager クライアントからエクスポートすることにより、入力パラメーター・タイプを定義しているスキーマが生成された場合、入力パラメーターのデータ・モデルは、階層型になる場合があります。階層型である場合、IBM Analytical Decision Management は、実行時にルール・サービスで必要なペイロードを直接作成しません。ペイロードの主な部分は、分析データ・ビュー・ソースによって生成されます。これにより、エキスパート・ユーザーが入力パラメーター・タイプを編集し、互換性を維持する方法が制限されます (例えば、エキスパート・ユーザーは、フィールド/要素がオプションであることを示すことしかできません)。

一般情報

IBM Operational Decision Management などのビジネス・ルール管理システムで作成されたルールを呼び出す前に、IBM Analytical Decision Management は、入力パラメーターのデータ・モデルが階層型とフラットのどちらであるのかをルールのサービスに尋ねます。入力パラメーターがフラットの場合は、ルールのサービスで必要な入力も検査されます。

入力パラメーターのデータ・モデルがフラットであり、スコアリングされるレコードでルール・サービスの必須フィールドのいずれかに Null 値が入っている場合、IBM Analytical Decision Management はルールを呼び出さず、ルールの出力/回答フィールドにはデフォルト値が設定されます。この値は、以下のように出力/回答フィールドのデータ型によって異なります。

表 6. 回答フィールドのデータ型の値 :

回答フィールドのデータ型	値
xsd:string	""
xsd:boolean	""
xsd:int	0
xsd:long	0
xsd:byte	0
xsd:double	0.0
xsd:float	0.0
xsd:date	0.0

表 6. 回答フィールドのデータ型の値 (続き):

回答フィールドのデータ型	値
xsd:time	0.0
xsd:datetime	0.0

これは、ルール・サービスを介したレコードのスコアリング結果に影響する可能性があるため、*.XSD* スキーマで入力パラメーターを編集する際には注意が必要です。

入力パラメーターのデータ・モデルが階層型であり、スコアリングされるレコードでルール・サービスの必須フィールドのいずれかに Null 値が入っている場合、IBM Analytical Decision Management はルールを呼び出しますが、Null 値が入っている 1 つ以上のフィールドが、データ・ビュー・ソース・ノードで作成されたペイロードから省略されます。省略されたフィールドが必須フィールドだった場合は、それが原因でルール呼び出しが失敗します。そのため、*.XSD* スキーマの入力パラメーターを編集するときは注意が必要です。

スキーマを編集するには、作成されるルール・サービスの知識が必要です。特定の入力フィールドがルールで不要であり、ルールの入力パラメーターのデータ・モデルがフラットの場合は、フィールド関連要素がスキーマの入力パラメーターから削除されるか、またはオプションとしてマークされる場合があります。特定のフィールドがルールで不要であり、ルールの入力パラメーターが階層型の場合は、フィールド関連要素を削除できません。ただし、オプションとしてマークすることは可能です。これは、IBM Analytical Decision Management は、階層データ・モデルに基づいてルールのペイロードを直接生成しないからです。

入力パラメーター・フィールドを削除すると、ルールが呼び出されるときにフィールド値が無視され、ルールはそのフィールドを参照できません。入力パラメーター・フィールドをオプションにすると Null 値が許容されるため、Null 値をチェックするための防御策をルールに記述する必要があります。例えば、ルールの入力パラメーターのデータ・モデルがフラットであり、Age フィールドがルールで使用されないことがわかっている場合は、関連要素を削除して、それが必須フィールドでないことを示すことができます。以下に示す *.XSD* スキーマの例では、フィールド Age が必要でなければ、`<xsd:element name="Age">` セクション全体をスキーマから削除できます。

```

<xsd:complexType name="drugs">
<xsd:annotation>
  <xsd:appinfo>
    <dmname>drugs</dmname>
  </xsd:appinfo>
</xsd:annotation>
<xsd:sequence>
  <xsd:element name="Age">
    <xsd:annotation>
      <xsd:appinfo>
        <dmname>Age</dmname>
      </xsd:appinfo>
    </xsd:annotation>
    <xsd:simpleType>
      <xsd:restriction base="xsd:long">
        <xsd:minInclusive value="15"/>
        <xsd:maxInclusive value="74"/>
      </xsd:restriction>
    </xsd:simpleType>
  </xsd:element>
  <xsd:element name="Sex">
    <xsd:annotation>
      <xsd:appinfo>
        <dmname>Sex</dmname>
      </xsd:appinfo>
    </xsd:annotation>

```

```

<xsd:simpleType>
  <xsd:restriction base="xsd:string">
    <xsd:enumeration value="F"/>
    <xsd:enumeration value="M"/>
  </xsd:restriction>
</xsd:simpleType>
</xsd:element>
...
</xsd:sequence>
</xsd:complexType>

```

あるいは、ルールの入力パラメーターのデータ・モデルがフラットまたは階層型の場合は、ルールで必要のない入力フィールド要素の `minOccurs` 属性値を 0 に設定して、その要素をオプションにすることができます。

```

<xsd:complexType name="drugs">
  <xsd:annotation>
    <xsd:appinfo>
      <dmname>drugs</dmname>
    </xsd:appinfo>
  </xsd:annotation>
  <xsd:sequence>
    <xsd:element name="Age" minOccurs="0">
      <xsd:annotation>
        <xsd:appinfo>
          <dmname>Age</dmname>
        </xsd:appinfo>
      </xsd:annotation>
      <xsd:simpleType>
        <xsd:restriction base="xsd:long">
          <xsd:minInclusive value="15"/>
          <xsd:maxInclusive value="74"/>
        </xsd:restriction>
      </xsd:simpleType>
    </xsd:element>
    <xsd:element name="Sex">
      <xsd:annotation>
        <xsd:appinfo>
          <dmname>Sex</dmname>
        </xsd:appinfo>
      </xsd:annotation>
      <xsd:simpleType>
        <xsd:restriction base="xsd:string">
          <xsd:enumeration value="F"/>
          <xsd:enumeration value="M"/>
        </xsd:restriction>
      </xsd:simpleType>
    </xsd:element>
    ...
  </xsd:sequence>
</xsd:complexType>

```

階層データ・モデルの例

```

<xsd:complexType name="Customer">
  <xsd:sequence>
    <xsd:element name="FirstName" type="xsd:string" minOccurs="0"/>
    <xsd:element name="Surname" type="xsd:string" minOccurs="0"/>
    <xsd:element name="Age" type="xsd:long"/>
    <xsd:element name="Sex" type="xsd:string"/>
    <xsd:element name="CustID" type="xsd:string"/>
    <xsd:element maxOccurs="unbounded" minOccurs="0" name="Orders" type="xom:Order"/>
  </xsd:sequence>
</xsd:complexType>
<xsd:complexType name="Order">
  <xsd:sequence>
    <xsd:element name="CustID" type="xsd:string"/>
    <xsd:element name="NoOfItems" type="xsd:long"/>
  </xsd:sequence>
</xsd:complexType>

```

```

<xsd:element name="TotalCost" type="xsd:long" minOccurs="0"/>
<xsd:element name="Date" type="xsd:date"/>
</xsd:sequence>
</xsd:complexType>

```

Null 値を処理するための防御策をルールに記述することもできます。この場合、スキーマでフィールドを optional とマークできます。以下に示す IBM Operational Decision Management ルールの例 *currentOfferIsReplacePhone* では、Null 値の処理が記述されています。

```

if
  the CURRENT OFFER of 'the customer' is not null
  and the CURRENT OFFER of 'the customer' is "Replace Phone"
then
  make it true that 'the response' is value ;
else
  make it false that 'the response' is value ;

```

回答パラメーターのデータ型の編集

IBM Analytical Decision Management にさまざまなデータ型を返す IBM Operational Decision Management ルールを作成し、利用することができます (選択で使用する *boolean*、集計ルールなどで使用する *numeric*、および *string*)。ただし、*string* 以外のデータ回答ルールを作成するには、回答タイプを含む XML スキーマ定義 (.XSD) を手動で編集する必要があります。

また、ルール・プロジェクトに、返すデータ型がそれぞれに異なる複数のルールを含めることもできます。そのためには、複数バージョンの .XSD を保持する必要があります (例えば、データ回答変数が *numeric* に設定されたバージョン、*string* の別バージョン、*boolean* の別バージョンなど)。

IBM Analytical Decision Management からダウンロードされた .XSD ファイルには、2 つの複合データ型が含まれています。1 つは、IBM Analytical Decision Management プロジェクト・データ・モデルを表し、IBM Operational Decision Management ルール・プロジェクトの入力パラメーターに使用されます。もう 1 つは、ルールで返されるデータ型を表し、IBM Operational Decision Management ルール・プロジェクトの出力パラメーターに使用されます。

IBM SPSS Collaboration and Deployment Services Deployment Manager からダウンロードした回答関連 .XSD ファイルには、回答タイプ定義が含まれており、1 つの複合データ型のみが含まれることになります。

上記の両方のケースで、返されるデータ型 (その名前には必ず「response」接尾辞が付加されます) は、デフォルトで *string* に設定される value 要素を定義します。戻り値の種類が異なるルール・プロジェクトを作成する場合は、IBM Operational Decision Management ルール・プロジェクトをインポートする前に、返されるデータ型の value 要素が要求どおりに定義されるように、ダウンロードした .XSD を手動で編集する必要があります。以下に例を示します。

```

<xsd:complexType name="claim_dataresponse">
  <xsd:annotation>
    <xsd:appinfo>
      <dmname>claim_dataresponse</dmname>
    </xsd:appinfo>
  </xsd:annotation>
  <xsd:sequence>
    <xsd:element name="value" type="xsd:string"/>
  </xsd:sequence>
</xsd:complexType>

```

```

<xsd:complexType name="claim_dataresponse">
  <xsd:annotation>
    <xsd:appinfo>
      <dmname>claim_dataresponse</dmname>

```

```
        </xsd:appinfo>
    </xsd:annotation>
    <xsd:sequence>
        <xsd:element name="value" type="xsd:boolean"/>
    </xsd:sequence>
</xsd:complexType>
```

```
<xsd:complexType name="claim_dataresponse">
    <xsd:annotation>
        <xsd:appinfo>
            <dmname>claim_dataresponse</dmname>
        </xsd:appinfo>
    </xsd:annotation>
    <xsd:sequence>
        <xsd:element name="value" type="xsd:long"/>
    </xsd:sequence>
</xsd:complexType>
```

第 9 章 ヘルプの入手

以下の 3 つのレベルのユーザー支援が提供されます。

- カスタム・ヘルプ。IBM Analytical Decision Management にパッケージされている既製アプリケーションごとに用意されています。ヘルプを起動するには、任意の画面またはダイアログでヘルプ・アイコンをクリックします。



- コーチ・テキスト。小さなポップアップ・ウィンドウに表示され、現在のアプリケーションまたはビジネス上の課題に関する詳細が示されます。コーチ・テキストは、特定のサイトまたはアプリケーション用に簡単に構成できる、カスタマイズされたコンテキスト・ヘルプの追加レイヤーを提供します。コーチ・テキストは、ユーザー・インターフェースの全域で、小さなコーチ・テキスト・アイコンをクリックすると表示されます。



- 標準ツールチップ。ページ内またはダイアログ内の個別コントロール用に提供されます。ツールチップを表示するには、ボタン、リンク、または他のコントロールの上にマウス・ポインターを移動します。

利用可能な資料

IBM Analytical Decision Management の資料は、さまざまな形式で提供されています。アプリケーションのヘルプにアクセスすると、オンラインの IBM Analytical Decision Management インフォメーション・センターが起動され、関連するヘルプ・トピックが表示されます。ヘルプがうまく起動しないときは、外部のインフォメーション・センター URL にアクセスする権限がない場合があるため、管理者に連絡してください。

インフォメーション・センターは、IBM Analytical Decision Management のすべての資料をオンライン上の 1 箇所に集めたものです。製品が稼働していないときでも常に使用可能です。インフォメーション・センターは、ブラウザのブックマークに登録することをお勧めします。

インターネットに接続していないとき、またはインフォメーション・センターが保守のために一時的に使用不可のときでも、これらの資料は Portable Document Format (*.pdf) 形式で参照することができます。PDF ファイルは、インストール・メディアに収録されていますが、IBM SPSS Collaboration and Deployment Services インストール・ディレクトリー (C:\Program Files\IBM\SPSS\Deployment\6.0\Server\documentation\DecisionManagement など) にもインストールされています。

Web では、PDF 版の資料 (<http://www-01.ibm.com/support/docview.wss?uid=swg27038937>) およびリリース・ノート (<http://www-01.ibm.com/support/docview.wss?uid=swg27038946>) を参照できます。

付録 A. ストリームの共有

IBM Analytical Decision Management で作成されたモデルとプロジェクトは、ストリーム・ファイルとして保存され、IBM SPSS Modeler が使用可能であれば、そこで変更し、使用することができます。例えば、IBM SPSS Modeler Advantage または別の IBM Analytical Decision Management アプリケーションを使用してモデルまたはプロジェクト・ストリームを作成し、そのストリームを IBM Analytical Decision Management を通じて展開する前に IBM SPSS Modeler でカスタマイズできます。あるいは、この付録で説明するように展開設定が指定されていれば、IBM SPSS Modeler で作成されたストリームを IBM Analytical Decision Management で開くことができます。

この付録で説明する形式に準拠したモデルは、IBM SPSS Modeler Advantage で表示、編集、スコアリング、およびリフレッシュすることができます。それ以外のモデルも IBM SPSS Modeler Advantage でスコアリングできますが、正しく表示されない可能性があり、リフレッシュすることもできません。

IBM Analytical Decision Management でストリームを開く

IBM Analytical Decision Management または IBM SPSS Modeler Advantage でストリームを開く場合

- 展開タイプが指定されていない場合 (なし)、ストリームを開くことはできますが、「データ」タブしか使用できません。
- 展開タイプが「スコアリングのみ」である場合は、スコアリング枝の開始点が IBM SPSS Modeler Advantage または IBM Analytical Decision Management でサポートされるタイプの単一データ入力ノードであることを条件として、「スコア」タブも使用できます。つまり、ストリーム内の入力ノードが 1 つである (データ・セットがマージされていない) 限り、IBM SPSS Modeler でサポートされるすべてのソース・タイプが IBM SPSS Modeler Advantage および IBM Analytical Decision Management でもサポートされます。
- 展開タイプが「モデルのリフレッシュ」である場合は、モデル作成枝の開始点が IBM SPSS Modeler Advantage または IBM Analytical Decision Management でサポートされるタイプの単一データ入力ノードであること、および選択したモデル作成ノードが自動化モデル作成ノード (予測モデル作成でサポート)、ディジション・リスト・ノード (予測ルール・モデル作成でサポート)、クラスタリング・ノード (クラスター・モデル作成でサポート)、またはアソシエーション・ルール・ノード (アソシエーション・モデル作成でサポート) であることを条件として、「モデル作成」タブも使用できます。さらに、ターゲット・フィールド、およびモデル作成で使用するその他の情報を定義するデータ型ノードが存在していることが必要です。

IBM SPSS Modeler での展開設定の指定

IBM SPSS Modeler で、以下のように「ストリームのプロパティ」ダイアログ・ボックスで展開設定を指定します。

1. IBM SPSS Modeler のメニューから以下のように選択します。

「ツール」 > 「ストリームのプロパティ」 > 「展開」

2. 「展開タイプ」リストで、「スコアリングのみ」または「モデルのリフレッシュ」を必要に応じて選択します。
3. 必要に応じて、「スコアリング・ノード」、「モデル作成ノード」、および「モデル・ノード」を選択します。

4. 「**チェック**」を選択して必要なオプションがすべて指定されていることを確認するか、または「**保管**」を選択してストリームを IBM SPSS Collaboration and Deployment Services Repository に格納します。

IBM SPSS Modeler Advantage または IBM Analytical Decision Management で作成されたストリームについては、展開情報が自動的に設定されます。

IBM SPSS Modeler ストリームを IBM Analytical Decision Management で使用するためのヒント

- ストリームは、IBM SPSS Modeler を使用して最初から作成するよりも、IBM SPSS Modeler Advantage または IBM Analytical Decision Management を使用して作成した後に IBM SPSS Modeler で変更する方が簡単です。この方法で作成されたストリームは、IBM Analytical Decision Management で完全に機能するように構築されます。
- IBM SPSS Modeler Advantage で作成されるストリームでは、エキスパート・ユーザーがカスタマイズできる枝内の特定の場所にスーパーノードが含まれています。ストリームに対するカスタマイズや追加は、スーパーノード内で行うことを強くお勧めします。それによって、カスタマイズまたは追加が確実にストリームに保存され、IBM SPSS Modeler Advantage で認識されるようになります。
- 各入力ノードのすぐ下流に式ルールを含んだローカルルール・ノードがある場合は、そのフィールドが「追加フィールドの追加/編集」ダイアログに表示されます。ローカルルールは、IBM Analytical Decision Management または IBM SPSS Modeler Advantage でのみ作成でき、IBM SPSS Modeler では作成できないことに注意してください。
- モデル作成ノードの上流 (およびデータ型ノードの下流) にデータ区分ノードがある場合は、IBM SPSS Modeler Advantage で「自動的にデータを区分して、検証およびテストのためのデータ・ソース作成時にモデル評価を実行できるようにする」が有効となり、データ区分ノードが有効かどうかに応じて、このチェック・ボックスの状態が設定されます。
- データ型ノードの下流に自動データ準備 (ADP) ノードがある場合は、「自動的にデータをクリーンアップして準備し、モデル作成の信頼性を確保する」が有効となり、ADP ノードが有効かどうかに応じて、このチェック・ボックスの状態が設定されます。
- モデル作成ノードの上流にルール・ノードまたは条件抽出ノードがある場合は、これらのノードを使用して「使用する選択の指定」セクションの値が設定されます。
- ストリーム・パラメーターが定義されている場合 (「ストリームのプロパティ」ダイアログ・ボックスの「パラメーター」タブ) は、IBM SPSS Modeler Advantage の「モデル作成」タブと「スコア」タブに「パラメーター」リンクが表示されます。ユーザーは、このリンクをクリックして、パラメーターの値を指定できます。

IBM SPSS Modeler からの IBM SPSS Modeler Advantage の起動

IBM SPSS Modeler のスプラッシュ画面から、IBM SPSS Modeler Advantage でモデルを作成することを選択できます。

付録 B. アクセシビリティ

アクセシビリティ機能は、運動障害や視覚障害などの身体障害を持つユーザーが情報技術製品を快適に使用できるようにサポートします。IBM は、年齢や能力を問わず、すべての人が便利に使用できる製品の提供に努めています。本製品では、標準的な Web ブラウザー・ナビゲーション・キーを採用しています。

IBM Analytical Decision Management はブラウザ・ベースのアプリケーションであるため、たとえば、視覚障害者のような障害者向けのアクセシビリティ・オプションは、Web ブラウザーの設定によって制御されます。キーボード・ショートカット、スクリーン・リーダーなどのユーザー・インターフェースの操作について詳しくは、ご使用の Web ブラウザーの資料を参照してください。

ベンダー・ソフトウェア

IBM Analytical Decision Management では、IBM 使用許諾契約書で保証されていない何らかのベンダー・ソフトウェアを使用する必要があることがあります。これらの製品のアクセシビリティ機能について、IBM は一切の保証責任を負いません。ベンダー製品のアクセシビリティ情報については、それぞれのベンダーにお問い合わせください。

IBM とアクセシビリティ

アクセシビリティに関する IBM のコミットメントについて詳しくは、IBM Human Ability and Accessibility Center を参照してください。

特記事項

本書は米国 IBM が提供する製品およびサービスについて作成したものです。

本書に記載の製品、サービス、または機能が日本においては提供されていない場合があります。日本で利用可能な製品、サービス、および機能については、日本 IBM の営業担当員にお尋ねください。本書で IBM 製品、プログラム、またはサービスに言及していても、その IBM 製品、プログラム、またはサービスのみが使用可能であることを意味するものではありません。これらに代えて、IBM の知的所有権を侵害することのない、機能的に同等の製品、プログラム、またはサービスを使用することができます。ただし、IBM 以外の製品とプログラムの操作またはサービスの評価および検証は、お客様の責任で行っていただきます。

IBM は、本書に記載されている内容に関して特許権 (特許出願中のものを含む) を保有している場合があります。本書の提供は、お客様にこれらの特許権について実施権を許諾することを意味するものではありません。実施権についてのお問い合わせは、書面にて下記宛先にお送りください。

〒103-8510

東京都中央区日本橋箱崎町19番21号

日本アイ・ビー・エム株式会社

法務・知的財産

知的財産権ライセンス渉外

以下の保証は、国または地域の法律に沿わない場合は、適用されません。IBM およびその直接または間接の子会社は、本書を特定物として現存するままの状態を提供し、商品性の保証、特定目的適合性の保証および法律上の瑕疵担保責任を含むすべての明示もしくは黙示の保証責任を負わないものとします。国または地域によっては、法律の強行規定により、保証責任の制限が禁じられる場合、強行規定の制限を受けるものとします。

この情報には、技術的に不適切な記述や誤植を含む場合があります。本書は定期的に見直され、必要な変更は本書の次版に組み込まれます。IBM は予告なしに、随時、この文書に記載されている製品またはプログラムに対して、改良または変更を行うことがあります。

本書において IBM 以外の Web サイトに言及している場合がありますが、便宜のため記載しただけであり、決してそれらの Web サイトを推奨するものではありません。それらの Web サイトにある資料は、この IBM 製品の資料の一部ではありません。それらの Web サイトは、お客様の責任でご使用ください。

IBM は、お客様が提供するいかなる情報も、お客様に対してなんら義務も負うことのない、自ら適切と信ずる方法で、使用もしくは配布することができるものとします。

本プログラムのライセンス保持者で、(i) 独自に作成したプログラムとその他のプログラム (本プログラムを含む) との間での情報交換、および (ii) 交換された情報の相互利用を可能にすることを目的として、本プログラムに関する情報を必要とする方は、下記に連絡してください。

IBM Software Group

ATTN: Licensing

200 W. Madison St.

Chicago, IL; 60606

U.S.A.

本プログラムに関する上記の情報は、適切な使用条件の下で使用することができますが、有償の場合もあります。

本書で説明されているライセンス・プログラムまたはその他のライセンス資料は、IBM 所定のプログラム契約の契約条項、IBM プログラムのご使用条件、またはそれと同等の条項に基づいて、IBM より提供されます。

この文書に含まれるいかなるパフォーマンス・データも、管理環境下で決定されたものです。そのため、他の操作環境で得られた結果は、異なる可能性があります。一部の測定が、開発レベルのシステムで行われた可能性があります。その測定値が、一般に利用可能なシステムのものと同じである保証はありません。さらに、一部の測定値が、推定値である可能性があります。実際の結果は、異なる可能性があります。お客様は、お客様の特定の環境に適したデータを確かめる必要があります。

IBM 以外の製品に関する情報は、その製品の供給者、出版物、もしくはその他の公に利用可能なソースから入手したものです。IBM は、それらの製品のテストは行っておりません。したがって、他社製品に関する実行性、互換性、またはその他の要求については確認できません。IBM 以外の製品の性能に関する質問は、それらの製品の供給者をお願いします。

IBM の将来の方向または意向に関する記述については、予告なしに変更または撤回される場合があります、単に目標を示しているものです。

本書には、日常の業務処理で用いられるデータや報告書の例が含まれています。より具体性を与えるために、それらの例には、個人、企業、ブランド、あるいは製品などの名前が含まれている場合があります。これらの名称はすべて架空のものであり、名称や住所が類似する企業が実在しているとしても、それは偶然にすぎません。

この情報をソフトコピーでご覧になっている場合は、写真やカラーの図表は表示されない場合があります。

商標

IBM、IBM ロゴおよび ibm.com は、世界の多くの国で登録された International Business Machines Corporation の商標です。他の製品名およびサービス名等は、それぞれ IBM または各社の商標である場合があります。現時点での IBM の商標リストについては、<http://www.ibm.com/legal/copytrade.shtml> をご覧ください。

Adobe、Adobe ロゴ、PostScript、PostScript ロゴは、Adobe Systems Incorporated の米国およびその他の国における登録商標または商標です。

インテル、Intel、Intel ロゴ、Intel Inside、Intel Inside ロゴ、Centrino、Intel Centrino ロゴ、Celeron、Xeon、Intel SpeedStep、Itanium、および Pentium は、Intel Corporation または子会社の米国およびその他の国における商標または登録商標です。

Linux は、Linus Torvalds の米国およびその他の国における登録商標です。

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。

UNIX は The Open Group の米国およびその他の国における登録商標です。

Java およびすべての Java 関連の商標およびロゴは Oracle やその関連会社の米国およびその他の国における商標または登録商標です。

索引

日本語, 数字, 英字, 特殊文字の順に配列されています。なお, 濁音と半濁音は清音と同等に扱われています。

[ア行]

アイコン, IBM Cognos BI 40
アクセシビリティ 93
アソシエーション・モデリングの例 18
アソシエーション・モデル 58, 62
 テーブル形式データとトランザクション・データ 59
 テーブル形式のデータ 59
 トランザクション・データ 59
アプリケーション
 起動 11
 テスト 66
アプリケーション URL 21
アプリケーション起動ページ 21, 79
 アプリケーションの追加 21
 カスタマイズ 21
アプリケーションのテスト 66
アプリケーション・ショートカット 21, 79
アプリケーション・ショートカットの制御 79
アプリケーション・ホーム・ページ 22
一致グラフ 63
一般オプション 78
インタラクティブ・リスト・モデル 55, 61
エンティティ・キー 42
オンライン・ヘルプ 89

[カ行]

回答グラフ 63
外部ルール
 アプリケーションでの使用 50
概要
 管理 77
管理
 アプリケーション・ショートカット 79
 一般オプション 78
 概要 77
 スコアリング宛先オプション 80
 データ・オプション 79
 ホーム・ページ 79

管理の概要 77
キー 42
キーワード 25, 27
起動ページ 11, 21, 79
 アプリケーションの追加 21
 カスタマイズ 21
ギャラリー
 実行時間の長いジョブ 23
 未完了の作業 23
共有ルール 48
クラスター・モデリングの例 16
クラスターリング・モデル 57, 61
ゲイン・グラフ 13, 63
結合キー 42
言語 21
言語の変更 21
コーチ・テキスト 89
誤差の分布図 63

[サ行]

サンプル・アプリケーション 10
サンプル・アプリケーション用成果物 9
サンプル・アプリケーション用データ 10
サンプル・アプリケーション用ファイル 9, 10
サンプル・データ 10
サンプル・ファイル 9
式エディター 43
式マネージャー 43
時刻の形式 23
事前構築アプリケーション 10
シミュレーション
 デフォルトの日付 23
シミュレーション実施日 23
手動クラスター 58
ショートカット
 アプリケーション 21
除外ルール 47
新機能 1
スコアリング宛先オプション 80
スコアリング宛先オプションのロック 80
スコアリング要約レポート 75
スコアリング・フィールドのマップ 73
ストリーム
 アップロード 25
ストリームのアップロード 25
ストリーム・ファイルのアップロード 25
制御
 一般オプション 78

セグメント・ルール
 モデルでの使用 45
設定 23
説明 27
選択ルール 47
測定の尺度 32

[タ行]

ターゲット・フィールド
 モデルの作成 13, 53
ダイレクト・メールの例 9
注釈
 ルールへの追加 50
ツールチップ 89
追加フィールドおよびテーブルの編集 42
追加フィールドとテーブルの追加/編集 42
追加フィールドとテーブルの追加/編集/編集 42
通貨の形式 23
データ概要 32
データ型 32
データのプレビュー 31, 32
データベース・ソース 37
データ・オプション 79
データ・オプションのロック 79
データ・ソース
 式エディター 43
 測定の尺度 32
 追加フィールドとテーブルの追加/編集 42
 データ概要 32
 データベース 37
 データ・ビュー 38
 定義 29, 34
 テキスト・ベース 34
 入力フィールド 42
 ファイル・ソース 34
 フィールド値の編集 33
 フィールドのマップ 43
 プレビュー 31
Analytic Server 41
Excel 34
Hadoop 41
HDFS 41
IBM Cognos BI 39
IBM SPSS Statistics 34
SPSS Collaboration and Deployment Services Enterprise View 38
データ・ビュー 38

データ・ファイル 10
テキスト・ベースのデータ・ファイル 34
デモ用成果物 9

[ナ行]

入力フィールド
定義 42
マップ 43

[ハ行]

バージョンのラベル付け 26
バージョン・ラベル 26
ビジネス・ルール 45
日付の形式 23
ファイルのオープン 25
ファイルの保存 25, 26
フィールド
派生 43
マップ 43
フィールド値の編集 33
フィールドのマップ 43, 73
ブラウザ言語の変更 21
プロジェクトのロック 27
プロジェクト・データ・ソース 29
プロジェクト・データ・モデル 29
プロパティ
キーワード 27
説明 27
プロジェクト 27
名義型フィールドの最大メンバー数 27
リポジトリ 27
分布図 54, 63
ヘルプ 89
ヘルプの入手 89
利用可能な資料 89
変数の重要度グラフ 54
ホーム・ページ 22, 79
ホーム・ページの非表示設定 79
包含ルール 47

[マ行]

マイ・データ・ソース 29
未完了の項目
ギャラリー 23
名義型フィールドの最大メンバー数 27
モデル
アソシエーション 58
宛先データベースの選択 70
宛先ファイルの選択 70
インタラクティブ・リスト 55
オプション設定 60, 61, 62

モデル (続き)
クラスタリング 57
手動クラスター 58
グラフ 54
結果 54
作成 13, 53
作成時間の制限 23
出力フィールドの選択 74
スコアの宛先の選択 70
スコアリング 16, 69, 74, 75
スコアリング要約レポート 75
スコアリング・フィールドのマップ 73
設定の指定 23
ターゲット・フィールド 13, 53
データ選択 69
テスト 15, 66
評価 13, 63
予測 53
予測ルール 55
利益のシミュレート 65
ルールでの使用 45
レコード選択 70
Analytic Server 宛先の選択 72
Cognos BI サーバー宛先の選択 72
モデルでの Cognos BI サーバー宛先 72
モデルでの宛先データベース 70
モデルでの宛先ファイル 70
モデルでの出力フィールド 74
モデルでのスコアの宛先 70
モデルでのデータ選択 69
モデルでの利益シミュレーション 65
モデルでの利益のシミュレーション 65
モデルでのレコード選択 70
モデルの Analytic Server 宛先 72
モデルのグラフ 54
モデルの作成
ターゲット・フィールド 13
モデルのスコアリング 16, 69, 74
宛先データベース 70
宛先ファイル 70
出力フィールドの選択 74
スコアリングの宛先の選択 70
データの選択 69
フィールドのマップ 73
Analytic Server 宛先 72
Cognos BI サーバー宛先 72
モデルのテスト 15, 66
モデルの評価 13, 63
インタラクティブ・リスト・モデル 15
クラスタリング・モデル 16, 18
予測ルール・モデル 15

[ヤ行]

ユーザー設定 23
予測の重要度グラフ 54
予測モデル 53
予測ルール・モデル 55, 61

[ラ行]

利益
最大化 9
シミュレート 13
利益グラフ 13, 63
リフト・グラフ 63
利用可能な資料 89
リンク 42
ルール 45
エクスポート 48
外部 50
共有 48, 50
再使用 48, 50
作成 45
セグメント 45
選択 47
選択ロジック 47
注釈 50
モデルでの使用 45
IBM Operational Decision Management 50
ルールのエクスポート 48

A

Analytic Server データ 41

C

Cognos BI ソース 39

E

Excel データ・ファイル 34

H

Hadoop データ 41

HDFS データ 41

I

IBM Analytical Decision Management と
IBM SPSS Modeler の間でのストリーム
の共有 91

IBM Cognos BI ソース 39

IBM Cognos BI 入力ノード
アイコン 40

IBM Operational Decision Management の
ルール
アプリケーションでの使用 50

IBM SPSS Collaboration and Deployment
Services Enterprise View データ・ソース
38

IBM SPSS Collaboration and Deployment
Services Repository
保存 25, 26

IBM SPSS Modeler Advantage
例 9

IBM SPSS Statistics データ・ファイル
34

O

Operational Decision Management
ルール 50

R

ROI グラフ 63

U

URL
アプリケーション 21



Printed in Japan