

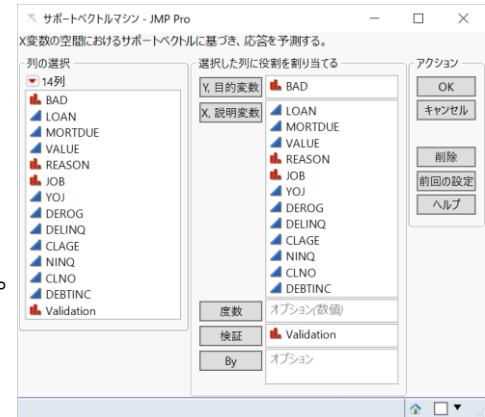
サポートベクトルマシン – 分類

JMP PRO 境界に基づいた統計モデルを使い、複数の予測変数の関数として、質的変数の応答を予測(分類)します。サポートベクトルマシンを用いると、分類木(パーティション)や判別分析よりもはるかに柔軟に境界の形を定めることができます。

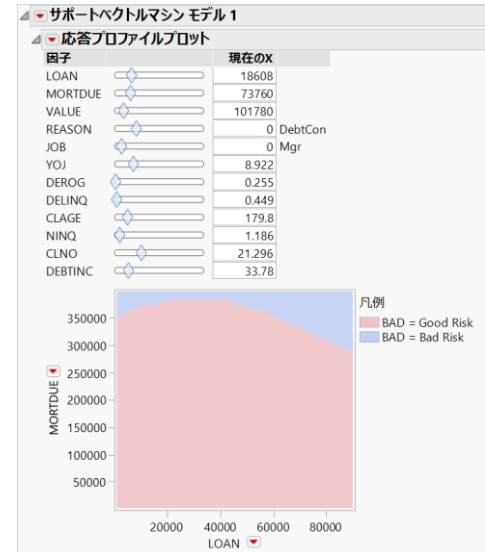
サポートベクトルマシン

例: Equity.jmp (ヘルプ > サンプルデータライブラリ)

1. データテーブルから、分析 > 予測モデル > サポートベクトルマシンを選択します。
2. 列の選択から名義尺度または順序尺度の応答変数を選択し、Y, 目的変数を選択します。
3. 候補予測変数を選択し、X, 説明変数をクリックします。
4. 必要な場合、右図のように検証列を選択して検証をクリックします。
5. OK をクリックします。モデルの設定パネルが開き、カーネル関数や関連するオプションを選択することができます。この例ではデフォルトの設定を用います。実行をクリックすると、JMP は以下の内容を表示します:



- 応答プロファイルプロットは二つの予測変数に対して分類領域を表示します。点を表示する場合は応答プロファイルプロットの横の赤い三角ボタンから選択します。プロットの予測変数は、各軸の変数名の横の赤い三角ボタンを選択して変更することができます。残りの予測変数の水準は、プロットの上のスライダーを用いて変更することができます。
- モデルの要約(未掲載)と分類性能を示した混同行列。



解釈:

- 検証データは 649 例。45 例 (6.9%) が誤分類された。45/(45+18) = 71% の Bad Risk のオブザベーションが Good Risk と誤分類された。また 0/(586+0) = 0% の Good Risk のオブザベーションが Bad Risk と誤分類された。

注意: この混同行列では各オブザベーションは予測確率が 0.0001 を上回るときに Good Risk と分類されています。一方、ある種の誤分類を減らすために別の閾値を選択するのが有利な場合もあります。例えば Bad Risk のオブザベーションが Good Risk と誤分類されるのを減らすために別の閾値を選択することが考えられます。

確率に対する閾値の設定
確率の閾値: 0.000107

学習				検証			
実測値	予測値	割合	誤分類率	実測値	予測値	割合	誤分類率
BAD	Good Risk	Bad Risk	0.0550	BAD	Good Risk	Bad Risk	0.0693
Good Risk	1.000	0.000		Good Risk	1.000	0.000	
Bad Risk	0.621	0.379		Bad Risk	0.714	0.286	
実測値	予測値 度数			実測値	予測値 度数		
BAD	Good Risk	Bad Risk		BAD	Good Risk	Bad Risk	
Good Risk	1871	0		Good Risk	586	0	
Bad Risk	113	69		Bad Risk	45	18	

注意:

ROC 曲線、リフトチャート、プロファイル、予測値の保存、予測式の保存、確率の保存、確率の計算式の発行等の追加のオプションを使うには一番上の赤い三角ボタンをクリックします。サポートベクトルマシンの追加情報に関しては、JMP のヘルプか予測モデルおよび発展的なモデル(ヘルプ > JMP ドキュメンテーションライブラリ以下)で「サポートベクトルマシン」と検索してご確認ください。