

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4940407号
(P4940407)

(45) 発行日 平成24年5月30日(2012.5.30)

(24) 登録日 平成24年3月9日(2012.3.9)

(51) Int.Cl.

F I

G06N 5/04 (2006.01)

G06N 5/04 580A

G06N 5/04 550J

請求項の数 7 (全 36 頁)

(21) 出願番号	特願2008-123070 (P2008-123070)	(73) 特許権者	301021533
(22) 出願日	平成20年5月9日(2008.5.9)		独立行政法人産業技術総合研究所
(62) 分割の表示	特願2003-338878 (P2003-338878)		東京都千代田区霞が関1-3-1
	の分割	(73) 特許権者	506015546
原出願日	平成15年9月29日(2003.9.29)		モデライズ株式会社
(65) 公開番号	特開2008-226269 (P2008-226269A)		東京都江東区青海二丁目4番6
(43) 公開日	平成20年9月25日(2008.9.25)	(74) 代理人	230104019
審査請求日	平成20年6月6日(2008.6.6)		弁護士 大野 聖二
特許法第30条第1項適用 2003年7月29日 社 団法人 電子情報通信学会発行の「電子情報通信学会研 究報告 信学技法 Vol. 103 No. 228」に 発表		(74) 代理人	100106840
			弁理士 森田 耕司
		(74) 代理人	100115679
			弁理士 山田 勇毅
		(74) 代理人	100105991
			弁理士 田中 玲子
最終頁に続く			

(54) 【発明の名称】 モデル作成装置、情報分析装置、モデル作成方法、情報分析方法およびプログラム

(57) 【特許請求の範囲】

【請求項1】

複数の変数について、当該複数の変数がとった値の複数の組合せを記憶した元統計データに基づいて、一の変数がとる値が他の変数がとる値に依存する依存関係を複数の変数について表したベイジアンネットの暫定的なモデルを暫定モデルとして作成するベイジアンネット暫定モデル作成手段と、

前記元統計データに含まれていない変数であって、前記元統計データの複数の変数のうちのいずれかの変数がとる値に依存する値をとる隠れ説明変数の候補変数と、当該いずれかの変数とを含んだ隠れ説明変数提供用統計データを記憶した隠れ説明変数提供用統計データ記憶手段と、

前記ベイジアンネット暫定モデル作成手段にて作成された前記暫定モデルを第1の画面に表示し、前記第1の画面上で前記暫定モデルを構成する複数の変数のうちのいずれかに対する指示を受け付けたときに、当該指示された変数を隠れ説明変数によって説明したい説明対象変数として受け付ける説明対象変数受付手段と、

前記隠れ説明変数提供用統計データ記憶手段に記憶された隠れ説明変数の候補変数と当該隠れ説明変数の候補変数を採用するか否かを選択するチェックボックスとを第2の画面に表示し、前記チェックボックスにチェックがなされた隠れ説明変数の候補変数を前記説明対象変数に対する隠れ説明変数として受け付ける隠れ説明変数受付手段と、

選択された前記隠れ説明変数のとる値が親変数のとる値に依存する第1の依存関係を、前記隠れ説明変数提供用統計データ記憶手段に記憶された統計データに基づいて求め、当

該第 1 の依存関係と、前記説明対象変数のとる値が前記親変数のとる値に依存する第 2 の依存関係とに基づいて、前記説明対象変数のとる値が前記隠れ説明変数のとる値に依存する第 3 の依存関係を求める隠れ依存関係算出手段と、

前記第 1 の依存関係および前記第 3 の依存関係に基づいて、前記説明対象変数のとる値が前記隠れ説明変数のとる値に依存し、前記隠れ説明変数のとる値が前記親変数のとる値に依存する関係を表すように、前記隠れ説明変数を前記暫定モデルに組み込む組込手段と、

を備え、

前記説明対象変数受付手段は、前記組込手段にて隠れ説明変数を組み込んだベイジアンネットのモデルを新たな暫定モデルとして前記第 1 の画面に表示し、説明対象変数の選択を受け付け、説明対象変数の選択がなされないで確定が入力された時点で、表示中の暫定モデルを最終的なベイジアンネットのモデルとして完成させることを特徴とするモデル作成装置。

【請求項 2】

前記隠れ説明変数受付手段は、前記第 2 の画面に、前記説明対象変数の親変数と、当該親変数を前記隠れ説明変数の候補変数の親変数として採用するか否かを選択するチェックボックスとを表示し、前記チェックボックスにチェックがなされた親変数を前記隠れ説明変数に対する親変数として受け付ける請求項 1 に記載のモデル作成装置。

【請求項 3】

ベイジアンネットのモデルに基づいて確率的推論を行う情報分析装置であって、

請求項 1 または 2 に記載のモデル作成装置にて作成されたベイジアンネットのモデルを記憶したベイジアンネットモデル記憶手段と、

前記ベイジアンネットモデル記憶手段に記憶された前記ベイジアンネットのモデルに基づいて、観測された変数の状態から未観測の変数の状態を推論する推論手段と、

を備えることを特徴とする情報分析装置。

【請求項 4】

複数の変数について、当該複数の変数がとった値の複数の組合せを記憶した元統計データに基づいて、モデル作成装置がベイジアンネットのモデルを作成するベイジアンネットモデル作成方法であって、

前記モデル作成装置が、複数の変数について、当該複数の変数がとった値の複数の組合せを記憶した元統計データに基づいて、一の変数がとる値が他の変数がとる値に依存する依存関係を複数の変数について表したベイジアンネットの暫定的なモデルを暫定モデルとして作成するベイジアンネット暫定モデル作成ステップと、

前記モデル作成装置が、前記元統計データに含まれていない変数であって、前記元統計データの複数の変数のうちのいずれかの変数がとる値に依存する値をとる隠れ説明変数の候補変数と、当該いずれかの変数とを含んだ隠れ説明変数提供用統計データを隠れ説明変数提供用統計データ記憶手段に記憶する隠れ説明変数提供用統計データ記憶ステップと、

前記モデル作成装置が、前記ベイジアンネット暫定モデル作成ステップにて作成された前記暫定モデルを第 1 の画面に表示し、前記第 1 の画面上で前記暫定モデルを構成する複数の変数のうちのいずれかに対する指示を受け付けたときに、当該指示された変数を隠れ説明変数によって説明したい説明対象変数として受け付ける説明対象変数受付ステップと、

前記モデル作成装置が、前記隠れ説明変数提供用統計データ記憶手段に記憶された変数を隠れ説明変数の候補変数と当該隠れ説明変数の候補変数を採用するか否かを選択するチェックボックスとを表示した第 2 の画面を前記第 1 の画面とともに表示し、前記チェックボックスにチェックがなされた隠れ説明変数の候補変数を前記説明対象変数に対する隠れ説明変数として受け付ける隠れ説明変数受付ステップと、

前記モデル作成装置が、選択された前記隠れ説明変数のとる値が親変数のとる値に依存する第 1 の依存関係を、前記隠れ説明変数提供用統計データ記憶手段に記憶された統計データに基づいて求め、当該第 1 の依存関係と、前記説明対象変数のとる値が前記親変数の

10

20

30

40

50

とる値に依存する第2の依存関係とに基づいて、前記説明対象変数のとる値が前記隠れ説明変数のとる値に依存する第3の依存関係を求める隠れ依存関係算出ステップと、

前記モデル作成装置が、前記第1の依存関係および前記第3の依存関係に基づいて、前記説明対象変数のとる値が前記隠れ説明変数のとる値に依存し、前記隠れ説明変数のとる値が前記親変数のとる値に依存する関係を表すように、前記隠れ説明変数を前記暫定モデルに組み込む組込ステップと、

を備え、

前記説明対象変数受付ステップは、前記組込ステップにて隠れ説明変数を組み込んだベイジアンネットのモデルを新たな暫定モデルとして前記第1の画面に表示し、説明対象変数の選択を受け付け、説明対象変数の選択がなされないで確定が入力された時点で、表示中の暫定モデルを最終的なベイジアンネットのモデルとして完成させることを特徴とするモデル作成方法。

【請求項5】

複数の変数について、当該複数の変数がとった値の複数の組合せを記憶した元統計データに基づいて、ベイジアンネットのモデルを作成するためのプログラムであって、コンピュータに、

複数の変数について、当該複数の変数がとった値の複数の組合せを記憶した元統計データに基づいて、一の変数がとる値が他の変数がとる値に依存する依存関係を複数の変数について表したベイジアンネットの暫定的なモデルを暫定モデルとして作成するベイジアンネット暫定モデル作成ステップと、

前記元統計データに含まれていない変数であって、前記元統計データの複数の変数のうちのいずれかの変数がとる値に依存する値をとる隠れ説明変数の候補変数と、当該いずれかの変数とを含んだ隠れ説明変数提供用統計データを隠れ説明変数提供用統計データ記憶手段に記憶する隠れ説明変数提供用統計データ記憶ステップと、

前記ベイジアンネット暫定モデル作成ステップにて作成された前記暫定モデルを第1の画面に表示し、前記第1の画面上で前記暫定モデルを構成する複数の変数のうちのいずれかに対する指示を受け付けたときに、当該指示された変数を隠れ説明変数によって説明したい説明対象変数として受け付ける説明対象変数受付ステップと、

前記隠れ説明変数提供用統計データ記憶手段に記憶された隠れ説明変数の候補変数と当該隠れ説明変数の候補変数を採用するか否かを選択するチェックボックスとを表示した第2の画面を前記第1の画面とともに表示し、前記チェックボックスにチェックがなされた隠れ説明変数の候補変数を前記説明対象変数に対する隠れ説明変数として受け付ける隠れ説明変数受付ステップと、

選択された前記隠れ説明変数のとる値が親変数のとる値に依存する第1の依存関係を、前記隠れ説明変数提供用統計データ記憶手段に記憶された統計データに基づいて求め、当該第1の依存関係と、前記説明対象変数のとる値が前記親変数のとる値に依存する第2の依存関係とに基づいて、前記説明対象変数のとる値が前記隠れ説明変数のとる値に依存する第3の依存関係を求める隠れ依存関係算出ステップと、

前記第1の依存関係および前記第3の依存関係に基づいて、前記説明対象変数のとる値が前記隠れ説明変数のとる値に依存し、前記隠れ説明変数のとる値が前記親変数のとる値に依存する関係を表すように、前記隠れ説明変数を前記暫定モデルに組み込む組込ステップと、

を実行させ、

前記説明対象変数受付ステップは、前記組込ステップにて隠れ説明変数を組み込んだベイジアンネットのモデルを新たな暫定モデルとして前記第1の画面に表示し、説明対象変数の選択を受け付け、説明対象変数の選択がなされないで確定が入力された時点で、表示中の暫定モデルを最終的なベイジアンネットのモデルとして完成させることを特徴とするプログラム。

【請求項6】

前記隠れ説明変数受付ステップは、前記第2の画面に、前記説明対象変数の親変数と、

当該親変数を前記隠れ説明変数の候補変数の親変数として採用するか否かを選択するチェックボックスとを表示し、前記チェックボックスにチェックがなされた親変数を前記隠れ説明変数に対する親変数として受け付ける請求項 5 に記載のプログラム。

【請求項 7】

ベイジアンネットのモデルに基づいて確率的推論を情報分析装置にて行うためのプログラムであって、コンピュータに、

請求項 5 または 6 に記載されたプログラムにて作成されたベイジアンネットのモデルをベイジアンネットモデル記憶手段に記憶させるステップと、

前記情報分析装置が、前記ベイジアンネットモデル記憶手段に記憶された前記ベイジアンネットのモデルに基づいて、観測された変数の状態から未観測の変数の状態を推論する推論ステップと、

を実行させるプログラム。

10

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、確率的推論に適用されるベイジアンネットのモデルを作成する装置および方法と、作成されたモデルに基づいて情報を分析する装置および方法、ならびにプログラムに関する。

【背景技術】

20

【0002】

情報処理技術の適用領域が拡大し、さまざまな状況や多様なユーザに適応して動作することができる情報処理メカニズムが重要になってきている。その場合に従来のプログラミングのように事前に規定したとおりに処理を進める方式だけでは対応できないことも多く、不確実な情報(たとえば、真か偽か断定できないような予測や不完全な観測情報)のもとでもできるだけ正しく状況を理解し、ユーザの意図を汲んで適切な処理を行う知的情報処理の仕組みが必要になってきている。

【0003】

これまでの知的情報処理ではできるだけ情報を明示化した記号として扱うアプローチを中心に発達してきたとも言える。コンピュータに入力された情報は比較的早い段階で記号化され内部的には記号のまま推論などの処理を行う仕組みであり、これは限られた計算機能力を効率的に利用するための必然だったとも考えられる。このアプローチは論理プログラミング、知的工学、エキスパートシステムといった狭義の人工知能研究の流れに沿って発展してきた。しかし知識を記号的に記述するにあたって必要とされる記述量の増大と、完全な記述が実際問題としてかなり困難であるといった問題を抱えていた。

30

【0004】

一方、あいまいで不確実な情報を確率論に基づいて取り扱い、内部的にもそのまま計算するアプローチがある。対象となる変数を、そのすべてのとり得る実現値に対する確信度すなわち確率分布として計算し、推論を行う確率的推論もその一つである。人工知能分野においては、1970年代にDempster-Shaffer理論、エキスパートシステムMYCINなどに先駆的な研究例をみることができる。

40

【0005】

確率論的推論は事後確率を最大化するパターン識別器などではごく自然に用いられてきたものでもある。しかし、意思決定理論に基づいてシステムを制御したり、有用な知識で表現し、複雑な処理を行うためには、比較的複雑な構造を持った確率モデルが必要になる。このような構造をもつ確率モデルの一つが、変数間の依存関係や因果関係を有向グラフで表すベイジアンネットである(非特許文献1)。

【0006】

ベイジアンネットは、不確実な事象の起こる確率の予測や期待効用を最大化する意思決定、障害診断などに利用することのできる確率モデルである。最近、リスク管理や予測な

50

どの不確実性を含む領域の拡大から多方面で注目されている。また、確率推論アルゴリズムの進歩やソフトウェアの普及も進みつつある。

【 0 0 0 7 】

しかし、一方でベイジアンネットを導入する際の難しさもまだ残されている。ベイジアンネット自体は古くから知られ、確率推論を実行するソフトウェアも登場していたが、実際のモデルに適用するうえでは、適切なベイジアンネットモデルを構築することが容易ではないという問題があった。モデル構築の際には、変数、条件付確率、グラフ構造を適切に決定する必要があるが、これまでの実用化例の多くは、問題領域に精通したエキスパートの経験や知識を必要としている（非特許文献 2）。

【 0 0 0 8 】

また、適切なベイジアンネットモデルの構築が容易ではないという問題を解決するため、最近では、統計データからモデルを学習するアルゴリズムが提案されている。現在良く知られているベイジアンネットモデル構築のためのアルゴリズムは、1) 統計データに含まれる一の変数の子変数とし、2) その変数に対して親変数の候補を一つずつ加えてモデルを作り、3) 評価が高くなったときだけ親変数として採用する、というものである。手順 1) ~ 3) を統計データ内のすべての変数について行うことにより、統計データからモデルを作成することができる。

【非特許文献 1】本村陽一「ベイジアンネットによる確率的推論技術」、計測と制御、Vol. 42 No. 8 2003 社団法人計測自動制御学会、p649-654

【非特許文献 2】本村陽一「ベイジアンネットソフトウェア BayoNet」、計測と制御、Vol. 42 No. 8 2003 社団法人計測自動制御学会、P693-694

【発明の開示】

【発明が解決しようとする課題】

【 0 0 0 9 】

しかし、上記アルゴリズムに代表される従来のベイジアンネットのモデル構築アルゴリズムでは、モデルが表現できる依存関係がモデルの元になる統計データに含まれる変数間の依存関係に限られていた。これにより、統計データに現れない変数が関係する依存関係が成立しており、しかもそのような依存関係によって確率的推論を行うのに適したモデルを作成できるとしても、その依存関係をモデル化できなかった。

【 0 0 1 0 】

本発明は、上記課題を解決し、変数間に成立している隠れた依存関係を説明する変数を組み込んで適切なモデルを作成可能なモデル作成装置、モデル作成方法と、作成されたモデルを用いて情報分析を行う情報分析装置および情報分析方法と、これらを実行するプログラムを提供することを目的とする。

【課題を解決するための手段】

【 0 0 1 1 】

本発明のモデル作成装置は、モデルの元になる元統計データに含まれる複数の変数とそれらの条件付確率に基づく依存関係で構成されるベイジアンネットのモデルを作成する装置であって、前記モデル内の子変数と少なくとも一の親変数との組合せに関連づけられるべき新たな変数であって、前記元統計データには含まれていない変数を隠れ説明変数として受け付ける受付手段と、前記少なくとも一の親変数と前記子変数との依存関係および前記少なくとも一の親変数と前記隠れ説明変数との依存関係に基づいて、前記隠れ説明変数と前記子変数との依存関係を求める隠れ依存関係算出手段と、前記少なくとも一の親変数と前記隠れ説明変数と前記子変数との依存関係により前記少なくとも一の親変数、前記隠れ説明変数および前記子変数が順次接続されるように、前記隠れ説明変数を前記モデルに組み込む組込手段とを備える。

【 0 0 1 2 】

この構成により、モデル内の子変数を説明するために、その元統計データには存在しない変数を隠れ説明変数として組み込むことができる。従って、モデル内の変数を適切に説明する隠れ説明変数を選択してモデルに組み込むことにより、適切なモデルを作成できる

10

20

30

40

50

。適切な隠れ説明変数を導入することにより、隠れ説明変数によって説明される子変数がとる状態のばらつきが小さくなり、子変数の状態を容易に推論できるモデルを作成できる。ここで、隠れ依存関係算出手段によって隠れ依存関係を求める実現例の一つを説明する。まず、元統計データが記憶された元統計データ記憶手段から、親変数および子変数に関するデータを読み出す。受付手段により受け付けられてモデル作成装置の記憶領域に記憶された隠れ説明変数と親変数との依存関係と、元統計データから読み出した親変数および子変数のデータから求められる依存関係とに基づいて、隠れ説明変数と子変数との依存関係を算出する。また、組込手段の実現例の一つを説明する。モデル記憶手段には、それぞれの変数を識別子により記憶すると共に各変数に関連づけて各変数の親変数の識別子を親情報として記憶する。組込手段は、隠れ説明変数が接続された子変数の親情報として隠れ説明変数の識別子を記憶し、隠れ説明変数の親情報として接続先の親変数の識別子を記憶する。

10

【0013】

前記隠れ依存関係算出手段は、前記元統計データと異なる隠れ変数提供用統計データであって、前記少なくとも一の親変数および前記隠れ説明変数およびそれらの依存関係を含んだ統計データを参照して前記隠れ説明変数と前記子変数の依存関係を算出してもよい。

【0014】

このように親変数および親変数に依存関係を有する隠れ説明変数を含む別の統計データを導入することにより、隠れ変数と親変数との依存関係を求めることができる。例えば、隠れ変数提供用統計データから隠れ説明変数および親変数に関するデータを読み出し、読み出したデータに基づいて依存関係を求める。

20

【0015】

上記モデル作成装置は、前記元統計データに含まれず、前記隠れ説明変数提供用統計データに含まれ、かつ、前記隠れ説明変数提供用データ内で前記親変数と依存関係を有する複数の変数を前記隠れ説明変数の候補として出力する候補出力手段を備え、前記受付手段は、出力された複数の前記候補の中から前記隠れ説明変数の選択を受け付けてもよい。

【0016】

この構成により、出力された候補の中から隠れ説明変数を選択できるので、適切な隠れ説明変数を容易に見つけることができる。

【0017】

本発明の別の態様のモデル作成装置は、モデルの元になる元統計データに含まれる複数の変数とそれらの条件付確率に基づく依存関係を示すリンクとによって表現されるベイジアンネットのモデルを作成する装置であって、前記モデル内のある子変数に関連づけられるべき新たな変数であって、前記元統計データには含まれていない変数を隠れ説明変数として受け付ける受付手段と、前記隠れ説明変数を經由するように、前記子変数と前記隠れ説明変数、および前記子変数の少なくとも一の親変数と前記隠れ説明変数がリンクによって接続されたモデルを出力するモデル出力手段と、を備える。

30

【0018】

この構成により、モデル内のある子変数を説明するために、その元統計データには存在しない変数を隠れ説明変数として受け付け、受け付けた隠れ説明変数を子変数およびその親変数をリンクにより接続したモデルを出力できる。従って、モデル内の変数を適切に説明する隠れ説明変数を選択することにより、適切なモデルを作成できる。適切な隠れ説明変数を導入することにより、隠れ説明変数によって説明される子変数がとる状態のばらつきが小さくなり、子変数の状態を容易に推論できるモデルを作成できる。モデル出力手段は、モデルを記憶したモデル記憶手段からモデルを読み出し、読み出したモデルに隠れ説明変数を追加してモデルを出力することにより実現可能である。

40

【0019】

本発明の情報分析装置は、ベイジアンネットのモデルに基づいて確率的推論を行う情報分析装置であって、モデル内の子変数と少なくとも一の親変数との組合せに関連づけられるべき新たな変数であって前記モデルの元になる元統計データには含まれていない隠れ説

50

明変数を、前記少なくとも一の親変数、前記隠れ説明変数および前記子変数が順次接続されるように組み込んで作成されたベイジアンネットのモデルを記憶した記憶手段と、記憶された前記モデルに基づいて、観測された変数の状態から未観測の変数の状態を推論する推論手段とを備える。

【0020】

この構成によれば、記憶されたベイジアンネットのモデルは、元統計データに含まれない変数をモデル内の変数を説明するための隠れ説明変数として組み込んで作成されたモデルであるので、記憶されたモデルに基づいて観測された変数の状態から未観測の変数の状態を適切に推論することができる。

【0021】

本発明のモデル作成方法は、モデルの元になる元統計データに含まれる複数の変数とそれらの条件付確率に基づく依存関係で構成されるベイジアンネットのモデルを作成する方法であって、モデル記憶手段から読み出されたモデル内の子変数と少なくとも一の親変数との組合せに関連づけられるべき新たな変数であって、前記元統計データには含まれていない変数を隠れ説明変数として受け付け、コンピュータの記憶手段に記憶する受付ステップと、前記少なくとも一の親変数と前記子変数との依存関係および前記少なくとも一の親変数と前記隠れ説明変数との依存関係に基づいて、前記隠れ説明変数と前記子変数との依存関係を求める隠れ依存関係算出ステップと、前記少なくとも一の親変数と前記隠れ説明変数と前記子変数との依存関係により前記少なくとも一の親変数、前記隠れ説明変数および前記子変数が順次接続されるように、受付ステップで記憶された前記隠れ説明変数を前記モデルに組み込み、前記モデル記憶手段に記憶する組込ステップとを備える。

【0022】

本発明の別の態様のモデル作成方法は、モデルの元になる元統計データに含まれる複数の変数とそれらの条件付確率に基づく依存関係を示すリンクとによって表現されるベイジアンネットのモデルをコンピュータにより作成する方法であって、前記モデル内のある子変数に関連づけられるべき新たな変数であって、前記元統計データには含まれていない変数を隠れ説明変数としてコンピュータが受け付け、受け付けた隠れ説明変数をコンピュータの記憶手段に記憶する受付ステップと、前記隠れ説明変数を経由するように、前記子変数と前記隠れ説明変数、および前記子変数の少なくとも一の親変数と前記隠れ説明変数がリンクによって接続されたモデルを、受付ステップで記憶された隠れ説明変数、およびモデル記憶手段に記憶された子変数および少なくとも一の親変数を読み出して出力するモデル出力ステップとを備える。

【0023】

これにより、本発明のモデル作成装置と同様に、隠れ説明変数を組み込んだ適切なモデルを作成できるという効果がある。また、本発明のモデル作成装置と同様の特徴を本発明のモデル作成方法に適用することも可能である。

【0024】

本発明の情報分析方法は、ベイジアンネットのモデルに基づいて確率的推論を行う情報分析方法であって、モデル内の子変数と少なくとも一の親変数との組合せに関連づけられるべき新たな変数であって前記モデルの元になる元統計データには含まれていない隠れ説明変数を、前記少なくとも一の親変数、前記隠れ説明変数および前記子変数が順次接続されるように組み込んで作成されたベイジアンネットのモデルをモデル記憶手段に記憶する記憶ステップと、前記モデル記憶手段に記憶されたモデルに基づいて、観測された変数の状態から未観測の変数の状態を推論する推論ステップとを備える。

【0025】

これにより、本発明の情報分析装置と同様に、隠れ説明変数を組み込んで作成されたモデルに基づいて、適切に推論を行えるという効果がある。

【0026】

本発明のプログラムは、コンピュータにより、モデルの元になる元統計データに含まれる複数の変数とそれらの条件付確率に基づく依存関係で構成されるベイジアンネットのモ

10

20

30

40

50

デルを作成させるプログラムであって、モデル記憶手段に記憶されたモデル内の子変数と少なくとも一の親変数との組合せに関連づけられるべき新たな変数であって、前記元統計データには含まれていない変数を隠れ説明変数として受け付け、コンピュータの記憶手段に記憶する受付ステップと、前記少なくとも一の親変数と前記子変数との依存関係および前記少なくとも一の親変数と前記隠れ説明変数との依存関係に基づいて、前記隠れ説明変数と前記子変数との依存関係を求める隠れ依存関係算出ステップと、前記少なくとも一の親変数と前記隠れ説明変数と前記子変数との依存関係により前記少なくとも一の親変数、前記隠れ説明変数および前記子変数が順次接続されるように、受付ステップで記憶された前記隠れ説明変数を前記モデルに組み込み、前記モデル記憶手段に記憶する組込ステップとを前記コンピュータに実行させる。

10

【0027】

また、本発明の別の態様のプログラムは、コンピュータにより、モデルの元になる元統計データに含まれる複数の変数とそれらの条件付確率に基づく依存関係を示すリンクとによって表現されるベイジアンネットのモデルを作成させるプログラムであって、前記モデル内のある子変数に関連づけられるべき新たな変数であって、前記元統計データには含まれていない変数を隠れ説明変数として受け付け、受け付けた隠れ説明変数をコンピュータの記憶手段に記憶させる受付ステップと、前記隠れ説明変数を経由するように、前記子変数と前記隠れ説明変数、および前記子変数の少なくとも一の親変数と前記隠れ説明変数がリンクによって接続されたモデルを、受付ステップで記憶された隠れ説明変数、およびモデル記憶手段に記憶された子変数および少なくとも一の親変数を読み出して出力するモデル出力ステップとを前記コンピュータに実行させる。

20

【0028】

これらのプログラムをコンピュータで実行することにより、本発明のモデル作成装置と同様に、コンピュータにより、隠れ説明変数を組み込んだ適切なモデルを作成できるという効果がある。また、本発明のモデル作成装置と同様の特徴を本発明のプログラムに適用することも可能である。

【0029】

本発明の別の態様のプログラムは、コンピュータにより、ベイジアンネットのモデルに基づいて確率的推論を行う情報分析のためのプログラムであって、モデル内の子変数と少なくとも一の親変数との組合せに関連づけられるべき新たな変数であって前記モデルの元になる元統計データには含まれていない隠れ説明変数を、前記少なくとも一の親変数、前記隠れ説明変数および前記子変数が順次接続されるように組み込んで作成されたベイジアンネットのモデルをモデル記憶手段に記憶する記憶ステップと、前記モデル記憶手段に記憶されたモデルに基づいて、観測された変数の状態から未観測の変数の状態を推論する推論ステップとを前記コンピュータに実行させる。

30

【0030】

このプログラムをコンピュータで実行させることにより、本発明の情報分析装置と同様に、コンピュータによって、隠れ説明変数を組み込んで作成されたモデルに基づいて、適切に推論を行えるという効果がある。

【0031】

本発明の記録媒体は、上記のいずれかのプログラムを記録している。

40

【0032】

上記したモデル作成のためのプログラムを記録した記録媒体は、コンピュータに読み取らせて実行することにより、本発明のモデル作成装置と同様に、隠れ説明変数を組み込んだ適切なモデルを作成できるという効果がある。上記した情報分析のためのプログラムを記録した記録媒体は、コンピュータに読み取らせて実行することにより、本発明の情報分析装置と同様に、隠れ説明変数を組み込んで作成されたモデルに基づいて、適切に推論を行えるという効果がある。

【0033】

また、本発明の別の態様の記録媒体は、確率的推論に適用されるベイジアンネットのモ

50

デルを記録した記録媒体であって、複数の変数間の条件付確率に基づく依存関係を表すモデルであって、モデル内の子変数と少なくとも一の親変数との組合せに関連づけられるべき新たな変数であって前記モデルの元になる元統計データには含まれていない隠れ説明変数を、前記少なくとも一の親変数、前記隠れ説明変数および前記子変数が順次接続されるように組み込んで作成されたベイジアンネットのモデルを記録している。

【 0 0 3 4 】

この記録媒体に記録されたベイジアンネットのモデルは、隠れ説明変数の組み込みを受け付けて作成されたモデルである。従って、この記録媒体をコンピュータに読み取らせることにより、適切なベイジアンネットのモデルをコンピュータに記憶させ、そのモデルに基づく適切な推論が可能となる。

10

【発明の効果】

【 0 0 3 5 】

本発明は、データ内の子変数を説明するために、そのデータには存在しない変数を隠れ説明変数として組み込むことができるので、モデル内の子変数を適切に説明する隠れ説明変数を選択してモデルに組み込むことにより、隠れ説明変数によって説明される子変数と取る状態のばらつきが小さくなり、子変数の状態を容易に推論できるモデルを作成できるという効果を有する。

【発明を実施するための最良の形態】

【 0 0 3 6 】

以下、本発明の実施形態に係るモデル作成装置および情報分析装置について説明する。以下の説明では、情報分析装置は、顧客の購買履歴から購買予測を行う顧客分析装置を例として説明する。

20

【 0 0 3 7 】

図 1 は、実施形態のモデル作成装置 1 0 および顧客分析装置 3 0 の構成を示す図である。モデル作成装置 1 0 は、顧客の購買履歴の統計データに基づいてベイジアンネットのモデルを作成する機能を有する。ここでは、モデル作成装置 1 0 によって、ある店の過去の一定期間（例えば 1 ヶ月）における顧客の購買履歴から、ベイジアンネットのモデルを作成する例について説明する。顧客分析装置 3 0 は、モデル作成装置 1 0 により作成されたモデルに基づいて購買動向を分析する機能を有し、例えば新しい商品の購買予測を行う。なお、以下に説明する例は、本発明のモデル作成装置および情報分析装置の一例にすぎず、例えばインターネットのアクセス動向を分析するモデルの作成およびその分析など、様々な分野に適用可能である。

30

【 0 0 3 8 】

図 2 は、モデル作成装置 1 0 の構成を示す図である。モデル作成装置 1 0 は、顧客の購買履歴に関する統計データを記憶する履歴データ記憶部 1 2 と、事前知識に関する統計データを記憶した事前知識データ記憶部 1 4 と、履歴データおよび事前知識データに基づいてモデルを作成するモデル作成部 1 6 と、作成されたモデルを表示する表示部 1 8 と、隠れ説明変数を入力するための隠れ説明変数入力部 2 0 と、作成されたモデルを記憶するモデルデータ記憶部 2 2 とを備える。本実施形態においては、モデル作成装置 1 0 が履歴データおよび事前知識データを記憶することとしているが、モデル作成装置 1 0 が必ずしも履歴データ記憶部 1 2 および事前知識データ記憶部 1 4 を有しなくてもよい。例えば、モデル作成装置 1 0 をネットワークに接続し、モデル作成の際にネットワークを通じて必要な履歴データまたは事前知識データを取得する構成とすることもできる。

40

【 0 0 3 9 】

履歴データ記憶部 1 2 は、顧客の購買行動の結果を示す履歴データを記憶する機能を有する。図 3 は履歴データ記憶部 1 2 に記憶された履歴データの例を示す図である。この例では、ある商品を購入した顧客のデータおよび商品に関する情報が履歴として記憶されている。図 3 に見られるように、履歴データは、「顧客 ID」「年齢」「性別」「商品情報」の各変数を有するデータである。「顧客 ID」は顧客を特定するための識別子、「年齢」「性別」は顧客の属性に関する情報である。「商品」は、顧客 ID により特定される顧

50

客が購買した商品に関する情報である。この履歴データは、例えば次のようにして集めることができる。店で使用可能なサービスカードの会員を募集し、会員が商品を購入する際にはレジにてサービスカードを提示してもらうことにより、顧客が購買した商品の履歴を蓄積することができる。また、顧客の属性については、サービスカード入会時にアンケートを行うことにより取得できる。

【 0 0 4 0 】

事前知識データ記憶部 1 4 は、事前知識データを記憶する機能を有する。事前知識データとは、直接の分析対象である顧客から得られる購買行動に関するデータとは異なるデータである。事前知識データは、別のアンケート調査などにより得られた統計データでもよいし、統計データから得られた知識、例えば「何歳以上の人は車を持っている割合が何 % である」というような知識でもよい。

10

【 0 0 4 1 】

図 4 は、事前知識データ記憶部 1 4 に記憶された事前知識データの例を示す図である。図 4 に見られるように、事前知識データは、「調査対象者 ID」「年齢」「性別」「商品選択基準」の各変数を有するデータである。「調査対象者 ID」は調査対象者を特定するための識別子であり、「年齢」「性別」は調査対象者の属性に関する情報である。「商品選択基準」は、商品を購入する際に何を重視して選択するかについての情報である。例えば、品質、値段、ブランドなどの選択基準に関する情報である。この事前知識データは、例えば、分析対象の顧客とは異なる集団に対するアンケート調査等により取得することができる。

20

【 0 0 4 2 】

モデル作成部 1 6 は、履歴データ記憶部 1 2 に記憶された履歴データからベイジアンネットのモデルを作成する機能と、事前知識データ記憶部 1 4 に基づいて指定された隠れ説明変数をモデルに組み込む機能を有する。また、モデル作成部 1 6 は、隠れ説明変数により説明される変数（以下、「説明対象変数」という）の親変数と隠れ説明変数との依存関係、および隠れ説明変数と説明対象変数との依存関係を求める機能を有する。

【 0 0 4 3 】

モデル作成部 1 6 の詳しい説明に先立って、ベイジアンネットのモデルについて説明する。ベイジアンネットのモデルは複数の確率変数の間の定性的な依存関係をグラフ構造によって表し、個々の変数の間の定量的な関係を条件付確率で表した確率モデルである。変数間の依存関係は原因から結果となる変数の向きを有向リンクで図示する。例えば、変数 X が原因、変数 Y が結果という関係は「変数 X → 変数 Y」と表し、変数 X と変数 Y の依存関係を条件付確率で表す。このとき、リンクの元にある変数 X を親変数、リンクの先に来る変数 Y を子変数という。

30

【 0 0 4 4 】

次に、モデル作成部 1 6 について説明する。モデル作成部 1 6 は、次の手順によりモデルを作成する機能を有する。すなわち、1) 履歴データに含まれる変数のうちの1つの変数を子変数としたときの親変数の候補の変数を選択する、2) 子変数と親変数の候補の変数を含む部分モデルのうち、尤度に関するスコアの最も高い部分モデルを作成する、3) 統計データ内のそれぞれの変数を子変数として作成した部分モデルを接続してベイジアンネットのモデルを作成する。なお、モデル作成部 1 6 は、上記の手順に限らず、公知の手順によってモデルを作成してもよい。

40

【 0 0 4 5 】

また、モデル作成部 1 6 は、履歴データに含まれない隠れ説明変数をモデルに組み込む機能を有する。以下、図 5 から図 8 を参照しながら、モデル作成部 1 6 によってモデルに隠れ説明変数を組み込む手順について説明する。

【 0 0 4 6 】

図 5 は、履歴データから作成されたモデルの例を示す図である。図 5 に示すモデルは、履歴データに含まれる変数のみを含む。ここで、事前知識データに含まれる変数 H が、変数 C を説明するための隠れ説明変数として入力されたとする。隠れ説明変数の入力、次

50

に説明する隠れ説明変数入力部 20 によって行われる。

【0047】

事前知識データには、変数 C が含まれていないので、変数 C と変数 H との依存関係は事前知識データからは明らかではない。まず、モデル作成部 16 は、事前知識データに含まれる変数 A、変数 B、変数 H に基づいて、変数 C の親変数である変数 A、変数 B と入力された変数 H との依存関係を求める。

【0048】

図 6 は、変数 A、変数 B と変数 H との依存関係を示す図である。事前知識データに基づいて、変数 H が親変数として変数 A、変数 B を有することが求められたことが分かる。

【0049】

モデル作成部 16 は、変数 A、変数 B と変数 C との依存関係、および変数 A、変数 B と変数 H との依存関係に基づいて、変数 H と変数 C との依存関係を求める。例えば、統計学において公知の EM アルゴリズムにより、変数 H と変数 C の依存関係を求めることができる。これにより、図 7 に示すように変数 H と変数 C との依存関係が得られる。なお、ここで「変数 A、変数 B と変数 C との依存関係」とは、変数 C のとる値が変数 A、変数 B のとる値に依存する依存関係を意味する。同様に、「変数 A、変数 B と変数 H との依存関係」とは、変数 H のとる値が変数 A、変数 B のとる値に依存する依存関係を意味し、「変数 H と変数 C との依存関係」とは、変数 C のとる値が変数 H のとる値に依存する依存関係を意味する。モデル作成部 16 は、求めた依存関係に基づいて隠れ説明変数を図 5 に示すモデルに組み込む。これにより、図 8 に示すように変数 C の隠れ説明変数 H を組み込んだモデルを作成する。

【0050】

再び、図 2 を参照して、図 2 中の隠れ説明変数入力部 20 について説明する。隠れ説明変数入力部 20 は、隠れ説明変数の入力を受け付ける機能を有する。隠れ説明変数入力部 20 は、まず、モデル内の変数のうち隠れ説明変数により説明したい変数（説明対象変数）の指定を受け付け、次に、隠れ説明変数として用いる変数を受け付ける。説明対象変数の指定の受け付けは、例えば、表示されたモデル内の変数の選択により、選択された変数を説明対象変数として受け付ける。また、隠れ説明変数の受け付けは、例えば、隠れ説明変数として用いることのできる変数の候補を表示して、その中から選択させることにより隠れ説明変数を受け付ける。隠れ説明変数として用いることのできる変数の候補は、例えば、親変数と依存関係を有する変数を事前知識データベースから抽出することによって得ることができる。

【0051】

次に、顧客分析装置 30 について説明する。顧客分析装置 30 は、モデル作成装置 10 により作成されたモデルに基づいて商品の購買確率を予測する装置である。

【0052】

図 9 は、顧客分析装置 30 の構成を示す図である。顧客分析装置 30 は、モデル作成装置 10 により作成されたベイジアンネットのモデルを記憶するモデルデータ記憶部 32 と、商品に関する情報の入力を受け付ける商品情報入力部 34 と、モデルデータ記憶部 32 に記憶されたモデルに基づいて商品の購買確率を推論する推論部 36 と、推論部 36 により推論された購買確率を表示する表示部 38 とを備えている。

【0053】

モデルデータ記憶部 32 には、図 10 に示すベイジアンネットのモデルが記憶されている。なお、このモデルは、履歴データ記憶部 12 に記憶された履歴データ（図 3 参照）から作成されたモデルである。

【0054】

図 10 に示すモデルによれば、商品の購買行動は、商品情報という変数と、ブランド、品質、値段の 3 つの商品選択基準の変数に依存している。ここで、購買行動の変数は、購買行動を起こしたか、起こさないかの 2 つの値を有する変数である。購買行動を起こした割合が購買確率を表す。3 つの商品選択基準は、顧客の年齢および性別という変数に依存

10

20

30

40

50

している。すなわち、顧客の年齢および性別に応じて商品の選択基準が変化し、それによって購買確率も変化することとなる。ここで、3つの商品選択基準は、履歴データには含まれない変数であり、購買確率を説明するための隠れ説明変数として組み込まれた変数である。また、このモデルにおいては、顧客の年齢および性別には、履歴データから求められた確率分布が与えられている。例えば、履歴データを取得した集団において、男女の割合は3：7であるとか、年齢別にみると20代が15%、30代が20%・・・といった確率分布が与えられている。

【0055】

なお、このモデルにおいて、購買行動の変数は、履歴データに含まれないが、次のようにして変数として求めることができる。すなわち、履歴データは、商品を購入した顧客に関するデータであり、商品を購入しなかった顧客についてのデータは含まないが、例えば、他の商品売り上げなどから推定される総来店客数から、商品を購入した顧客の数を引くことにより、商品を購入しなかった人数を概算できる。この情報に基づいて、購買行動の変数がとるそれぞれの状態（商品を購入した、または購入しなかった）の確率を求めることができ、モデル内で変数として取り扱える。

【0056】

商品情報入力部34は、購買確率を予測したい商品に関する情報の入力を受け付ける機能を有する。ここで入力される情報としては、例えば商品の開発コンセプトである。開発コンセプトの例としては、品質を重視した商品、あるいは安さを追求した商品という情報がある。

【0057】

推論部36は、ベイジアンネットのモデルに基づいて、入力された商品に関する情報から購買確率を推論する機能を有する。

【0058】

ここで、ベイジアンネットに基づいて推論を行う方法について説明する。ベイジアンネットによる確率的推論は、1)観測された変数の値 e をノードにセットする、2)親ノードも観測値も持たないノードに部分確率を与える、3)知りたい対象の変数 X の事後確率 $P(X|e)$ を得る、という手順で行う。3)における事後確率を求めるために、観測された情報からの確率伝搬(変数間の局所計算)によって各変数の確率分布を更新していく確率伝搬法(belief propagation)と呼ばれる計算方法を用いる。

【0059】

表示部18は、推論結果を表示する機能を有する。これにより、分析者は、商品の売れ行きを予測することができる。

【0060】

次に、実施形態のモデル作成装置10の動作について説明する。モデル作成装置10によるモデル作成の動作を概説すると、まず、1)モデル作成部16により履歴データからベイジアンネットのモデルを作成する。続いて、2)作成されたモデルを表示部18に表示し、隠れ説明変数の入力を受け付け、入力された隠れ説明変数をモデルに組み込む、3)変更されたモデルを表示部18により表示し、隠れ説明変数の入力を受け付ける、という手順で作成される。なお、手順2)手順3)は、モデルが完成するまで繰り返される。

【0061】

図11は、モデル作成装置10の動作を示すフローチャートである。次に、図11を参照しながら実施形態に係るモデル作成装置10の動作について説明する。

【0062】

まず、モデル作成装置10は、履歴データからベイジアンネットのモデルを作成し、作成されたモデルを表示部18により表示する(S10)。

【0063】

図12は、作成したモデルを表示部18により表示した画面の例を示す図である。図12に示すように、図12に示す画面右側のモデル表示画面40に作成されたベイジアンネットのモデルが表示される。画面左側の隠れ説明変数入力画面42は、隠れ説明変数の入

10

20

30

40

50

力を受け付けるユーザインターフェースを提供している。

【0064】

次に、モデル作成装置10は、隠れ説明変数により説明したい変数の指定を受け付ける。ここでは、図12に示すモデル表示画面40に表示されたモデルにおいて、隠れ説明変数により説明する変数を指定する(S12)。例えば、マウスなどの操作により一の変数を示すノードを選択すると、その変数が隠れ説明変数により説明される説明対象変数として指定される。図12では、説明対象変数として購買行動の変数が選択されている。

【0065】

続いて、モデル作成装置10は、図12に示す画面左側の隠れ説明変数入力画面42において、隠れ説明変数の候補を表示する(S14)。ここでは、事前知識データに含まれる変数の一覧を「隠れ説明変数選択」ボックス44の「変数」欄50に表示している。

10

【0066】

この「隠れ説明変数選択」ボックス44において、「採否」欄46は、その候補の変数をモデル作成者が採用するか否かを選択するチェックボックスである。「親変数」欄48は、隠れ説明変数の親変数を選択するチェックボックスである。ここで選択される親変数は、隠れ説明変数と説明対象変数との依存関係を求めるために用いる親変数である。従って、この例では、「親変数」として指定可能な変数は、説明対象変数「購買行動」の親変数である「年齢」「性別」「商品情報」のいずれかである。

【0067】

図12に示す隠れ説明変数入力画面42の「隠れ説明変数選択」ボックス44において、モデル作成者が隠れ説明変数および親変数を指定してOKボタン52を選択すると、モデル作成装置10は隠れ説明変数の指定を受け付ける。隠れ説明変数の指定を受け付けると、モデル作成装置10は、親変数と隠れ説明変数との依存関係を求める(S18)。続いて、モデル作成装置10は、親変数と隠れ説明変数との依存関係、および親変数と説明対象変数との依存関係に基づいて、隠れ説明変数と説明対象変数との依存関係を求める(S20)。

20

【0068】

次に、モデル作成装置10は、親変数と隠れ説明変数との依存関係、および隠れ説明変数と説明対象変数との依存関係に基づいてモデルに隠れ説明変数を組み込み、表示部18に表示する(S22)。図13は、隠れ説明変数を組み込んだモデルを表示した画面を示す図である。

30

【0069】

次に、隠れ説明変数によって説明する変数が他にあるか否かを判定する(S24)。この判定は、モデル作成者によるモデル作成装置10への入力に基づいて行われる。すなわち、隠れ説明変数を組み込んだ後、図13に示す隠れ説明変数入力画面42において、モデル作成者が確定ボタン54を選択した場合には、隠れ説明変数によって説明すべき変数がないと判定する。図13に示す隠れ説明変数入力画面42において、モデル作成者が確定ボタン54を選択しないで、モデル表示画面40で次の説明対象変数を指定した場合には、隠れ説明変数によって説明すべき変数が他にあると判定する。

【0070】

40

隠れ説明変数によって説明すべき変数が他にあると判定された場合には、モデル作成装置10は、隠れ説明変数の候補の変数を表示するステップS14に遷移し、上記と同様にステップS14からステップS24までの処理を繰り返す。隠れ説明変数によって説明すべき変数が他にないと判定された場合には、モデル作成装置10は、作成したベジアンネットのモデルをモデルデータ記憶部22に記憶する(S26)。その後、モデルの作成の動作を終了する。

【0071】

図14は、顧客分析装置30の動作を示すフローチャートである。図14を参照しながら、顧客分析装置30の動作について説明する。

【0072】

50

まず、顧客分析装置 30 は、分析対象の商品の情報の入力を受け付ける (S30)。ここで受け付ける情報は、例えば、品質重視あるいは値段重視といった商品の開発コンセプトに関する情報である。

【0073】

次に、顧客分析装置 30 は、推論部 36 によってモデルデータ記憶部 22 に記憶されたモデルに基づいて購買確率の推論を行う (S32)。すなわち、モデル作成装置 10 は、履歴データにより確率分布が与えられている顧客の年齢および性別から、顧客がブランドで商品を選択する確率、品質で商品を選択する確率、値段で商品を選択する確率をそれぞれ求める。入力された商品の情報と、それぞれの選択基準の確率に基づいて顧客の購買行動の変数を推論する。購買行動を起こす人の割合が、商品の購買確率である。そして、顧客分析装置 30 は、求められた購買確率を表示部 38 により表示する (S34)。

10

【0074】

実施形態に係るモデル作成装置 10 は、履歴データにより作成されたモデル内の変数を説明するために、履歴データに含まれていない変数を隠れ説明変数として受け付け、受け付けた隠れ説明変数をモデルに組み込むことができる。従って、隠れ説明変数として適切な変数を組み込めば、適切なベイジアンネットのモデルを作成できる。説明対象変数がある状態のばらつきが少なくなるため、子変数の状態を容易に推論できるモデルを作成できる。

【0075】

実施形態に係る顧客分析装置 30 は、モデルに組み込む隠れ説明変数を受け付けて作成された適切なモデルに基づいて、購買確率の適切な推論を行うことができる。

20

【0076】

次に、本発明の実施形態に係るプログラムについて説明する。

【0077】

図 15 は、実施形態に係るモデル作成のためのプログラム 60 の構成を示す図である。プログラム 60 は、履歴データ取得モジュール 62 と、モデル作成モジュール 64 と、モデル表示モジュール 66 と、隠れ説明変数入力モジュール 68 と、モデルデータ記憶モジュール 70 とを有する。

【0078】

履歴データ取得モジュール 62 は、プログラム 60 をコンピュータで実行することにより、コンピュータに、履歴データを取得させるモジュールである。これにより、履歴データ記憶部とのインターフェースを実現できる。

30

【0079】

モデル作成モジュール 64 は、プログラム 60 をコンピュータで実行することにより、コンピュータに、モデルを作成させると共にモデルに隠れ説明変数を組み込ませるモジュールである。モデル作成モジュール 64 を実行することにより、上記実施形態のモデル作成装置 10 のモデル作成部 16 を実現できる。

【0080】

モデル表示モジュール 66 は、プログラム 60 をコンピュータで実行することにより、コンピュータに、作成されたモデルを表示させるモジュールである。モデル表示モジュール 66 を実行することにより、上記実施形態のモデル作成装置 10 の表示部 18 を実現できる。

40

【0081】

隠れ説明変数入力モジュール 68 は、プログラム 60 をコンピュータで実行することにより、コンピュータに、隠れ説明変数の入力を受け付けさせるモジュールである。隠れ説明変数入力モジュール 68 を実行することにより、上記実施形態のモデル作成装置 10 の隠れ説明変数入力部 20 を実現できる。

【0082】

モデルデータ記憶モジュール 70 は、プログラム 60 をコンピュータで実行することにより、コンピュータに、作成されたモデルを記憶させるモジュールである。モデル記憶モ

50

ジュールを実行することにより、モデルデータ記憶部とのインターフェースを実現できる。

【0083】

このモデル作成用のプログラム60をコンピュータで実行することにより、上記実施形態のモデル作成装置10を実現でき、隠れ説明変数を組み込むことにより適切なモデルを作成できるという上記実施形態のモデル作成装置10と同様の効果を有する。

【0084】

図16は、実施形態に係る顧客分析のためのプログラム80の構成を示す図である。プログラム80は、モデルデータ記憶モジュール82と、商品情報入力モジュール84と、推論モジュール86と、表示モジュール88とを有する。

10

【0085】

モデルデータ記憶モジュール82は、プログラム80をコンピュータで実行することにより、コンピュータに、ベイジアンネットのモデルを記憶させるモジュールである。モデルデータ記憶モジュール82によって記憶されるモデルは、隠れ説明変数を組み込んで作成された適切なモデルである。このモデルデータ記憶モジュール86により、コンピュータに上記実施形態の顧客分析装置30におけるモデルデータ記憶部を構成することができる。

【0086】

商品情報入力モジュール84は、プログラム80をコンピュータで実行することにより、コンピュータに、分析対象の商品に関する情報の入力を受け付けさせるモジュールである。商品情報入力モジュール84を実行することにより、上記実施形態の顧客分析装置30における商品情報入力部34を実現できる。

20

【0087】

推論モジュール86は、プログラム80をコンピュータで実行することにより、コンピュータに、ベイジアンネットのモデルに基づいて、商品に関する情報から購買確率を推論させるモジュールである。推論モジュール86を実行することにより、上記実施形態の顧客分析装置30における推論モジュール86を実現できる。

【0088】

表示モジュール88は、プログラム80をコンピュータで実行することにより、コンピュータに、推論された購買確率の結果を表示させるモジュールである。表示モジュール88を実行することにより、上記実施形態の顧客分析装置30における表示部38を実現できる。

30

【0089】

この顧客分析用のプログラム80をコンピュータで実行することにより、上記実施形態の顧客分析装置30を実現でき、隠れ説明変数を組み込んで作成された適切なモデルに基づいて商品の購買確率を推論できるという上記実施形態の顧客分析装置30と同様の効果を有する。

【0090】

以上、本発明のモデル作成装置10、顧客分析装置30、およびモデル作成方法、顧客分析方法ならびにプログラムについて、実施形態を挙げて説明したが、本発明は上記実施形態に限定されるものではない。

40

【0091】

上記実施形態では、隠れ説明変数の候補として、事前知識データベースに含まれる変数を隠れ説明変数入力画面42に表示し、モデル作成者に親変数と説明対象変数を選択させる構成としたが、図17に示すように、親変数と依存関係を有する変数のみを説明対象変数の候補として表示することとしてもよい。親変数と依存関係を有する変数のみを説明対象変数の候補として表示することにより、モデル作成者が親変数と依存関係のない変数を隠れ説明変数として選択しないようにできる。これにより、モデル作成装置は、親変数と隠れ説明変数、および親変数と説明対象変数の依存関係に基づいて隠れ説明変数と説明対象変数との依存関係を必ず求めることができ、選択された隠れ説明変数をモデルに組み込

50

むことができる。

【0092】

また、上記実施形態では、隠れ説明変数入力画面42におけるモデル作成者の変数選択に基づいてモデル作成装置10が親変数と隠れ説明変数との依存関係を求める構成について説明したが、モデル作成者が隠れ説明変数を設定し、親変数と隠れ説明変数との依存関係を入力することとしてもよい。

【0093】

図18は、隠れ説明変数入力画面42において、隠れ説明変数と親変数との依存関係を入力する例を示す図である。「依存関係入力」ボックス56において、年齢(親変数)に対応する可処分所得(隠れ説明変数)の入力を受け付ける構成となっている。この構成により、入力された依存関係に基づいて、隠れ説明変数と説明対象変数との依存関係を容易に求めることができる。

10

【0094】

なお、このようなインターフェースによって隠れ説明変数を入力する構成においては、事前知識データ記憶部14は必ずしも必要ではなく、またモデル作成部16も説明対象変数の親変数と隠れ説明変数との依存関係を求める機能は必ずしも必要ではない。

【0095】

また、上記した実施形態においては、モデル作成装置10と顧客分析装置30を別々に構成した例について説明したが、モデル作成装置10と顧客分析装置30を同じ装置によって構成してもよい。図19は、モデル作成と顧客分析の両方の機能を有する情報処理装置90の構成を示す図である。構成要素のそれぞれが有する機能は、上記実施形態のモデル作成装置10および顧客分析装置30の対応する構成要素のそれぞれと同じ機能を有する。この情報処理装置90により、顧客分析の分析者は、商品に応じて隠れ説明変数を変えてモデルを作成し、作成されたモデルを見て変数間の依存関係を分析することができる。これにより、商品に応じて何の変数が購買確率に影響するかを分析できる。

20

【0096】

また、上記した実施形態においては、モデル作成装置10はユーザインターフェース部分である表示部18および隠れ説明変数入力部20と、モデル作成部16とが一体に構成されていたが、ユーザインターフェースの部分とモデル作成部16とは必ずしも一体である必要はなく、別々の装置によって構成してもよい。図20は、モデル作成部16を有するモデル作成装置100と、モデル作成の指示を行う操作装置104とが別々に構成されている。モデル作成装置100と操作装置104はそれぞれ通信部102、106を有し、通信可能とされている。このような構成を採用することにより、ネットワーク上に設けられたモデル作成装置100をリモートで操作して、モデルの作成を行える。また、一台のモデル作成装置100に対して複数台の操作装置104を接続することにより、モデル作成装置100を共有できる。

30

【0097】

上記した実施形態において、顧客分析装置30は、入力された商品の情報に基づいて、その購買確率を求める例を説明したが、モデルデータ記憶部32に記憶されたモデルを用いて異なる顧客分析を行うこともできる。例えば、商品の購買確率を推論するのではなく、商品を購入する顧客の属性を推論してもよい。これにより、例えば、どのような顧客をターゲットに商品の宣伝を行えばよいかを知ることができる。

40

【産業上の利用可能性】

【0098】

以上説明したように、本発明は、統計データからベイジアンネットのモデルを作成する技術、および作成されたモデルに基づく推論を行う技術に関し、ベイジアンネットによる確率的推論の分野等において有用である。

【図面の簡単な説明】

【0099】

【図1】実施形態のモデル作成装置および顧客分析装置を示す図である。

50

- 【図 2】実施形態のモデル作成装置の構成を示す図である。
- 【図 3】履歴データの例を示す図である。
- 【図 4】事前知識データの例を示す図である。
- 【図 5】隠れ説明変数のモデルへの組み込み手順について説明する図である。
- 【図 6】隠れ説明変数のモデルへの組み込み手順について説明する図である。
- 【図 7】隠れ説明変数のモデルへの組み込み手順について説明する図である。
- 【図 8】隠れ説明変数のモデルへの組み込み手順について説明する図である。
- 【図 9】実施形態の顧客分析装置の構成を示す図である。
- 【図 10】顧客分析装置のモデルデータ記憶部に記憶されたモデルの例である。
- 【図 11】モデル作成装置の動作を示すフローチャートである。
- 【図 12】隠れ説明変数入力画面の例を示す図である。
- 【図 13】隠れ説明変数入力画面の例を示す図である。
- 【図 14】顧客分析装置の構成を示す図である。
- 【図 15】モデル作成のためのプログラムの構成を示す図である。
- 【図 16】顧客分析のためのプログラムの構成を示す図である。
- 【図 17】隠れ説明変数入力画面の他の例を示す図である。
- 【図 18】隠れ説明変数入力画面の他の例を示す図である。
- 【図 19】モデル作成と顧客分析の機能を有する情報処理装置の構成を示す図である。
- 【図 20】モデル作成装置の他の構成例を示す図である。

10

【符号の説明】

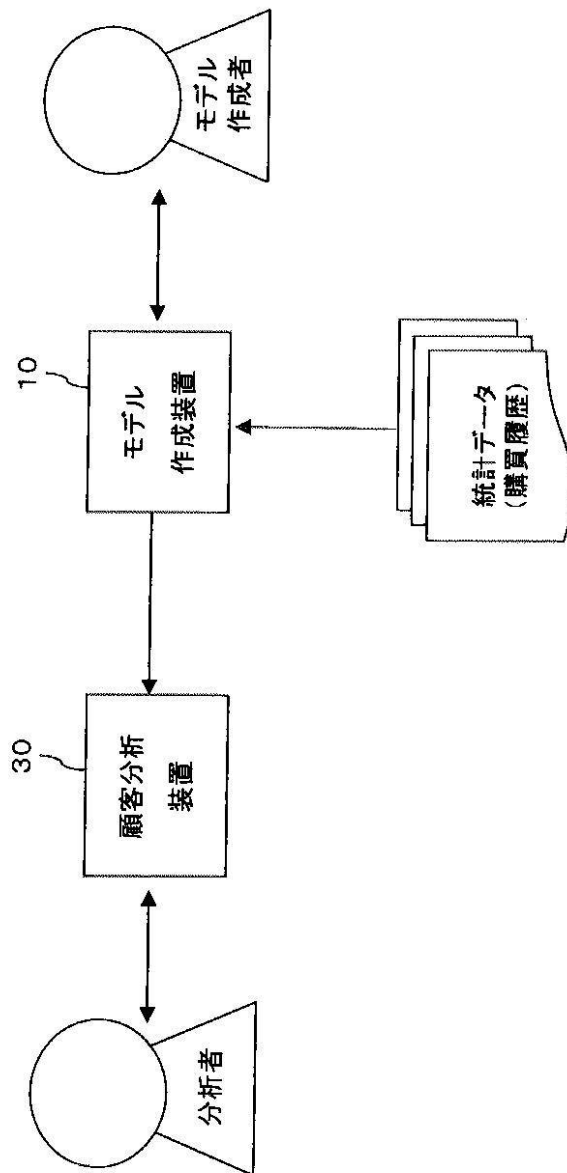
20

【 0 1 0 0 】

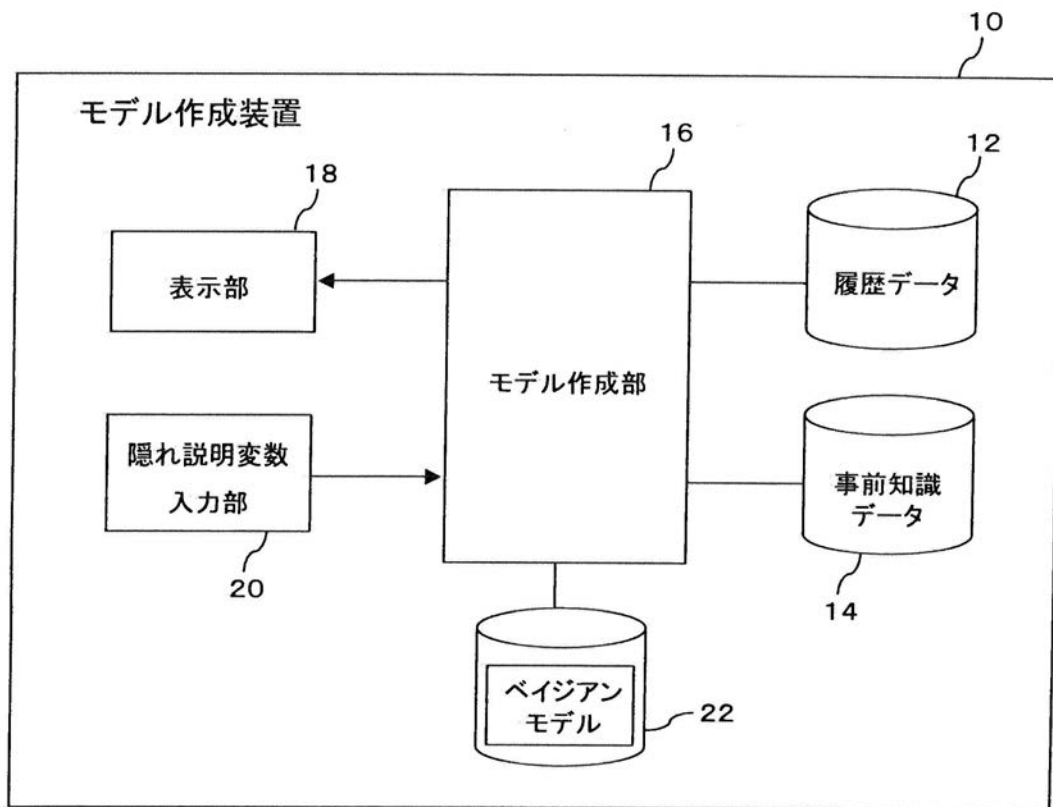
- 1 0 モデル作成装置
- 1 2 履歴データ記憶部
- 1 4 事前知識データ記憶部
- 1 6 モデル作成部
- 1 8 表示部
- 2 0 隠れ説明変数入力部
- 2 2 モデルデータ記憶部
- 3 0 顧客分析装置
- 3 2 モデルデータ記憶部
- 3 4 商品情報入力部
- 3 6 推論部
- 3 8 表示部
- 4 0 モデル表示画面
- 4 2 隠れ説明変数入力画面

30

【図 1】



【図 2】



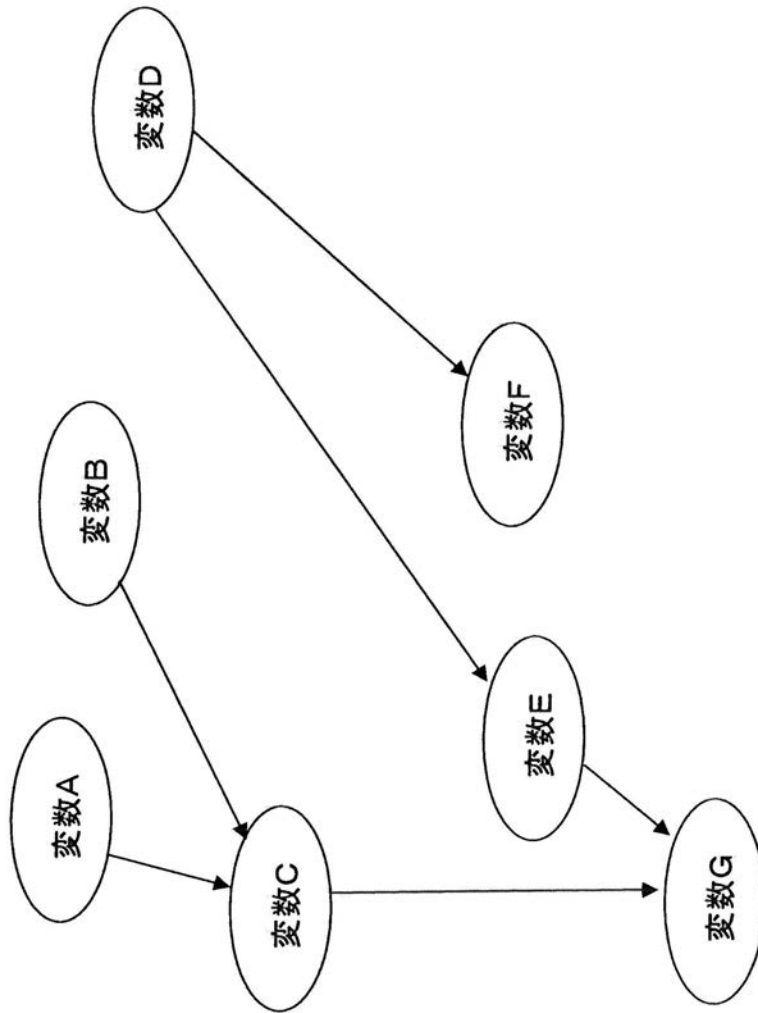
【図 3】

顧客ID	年齢	性別	商品情報
C10010	24	男	セール品
C10012	30	女	ブランド品
C10023	50	男	品質保証品
・	・	・	・
・	・	・	・
・	・	・	・

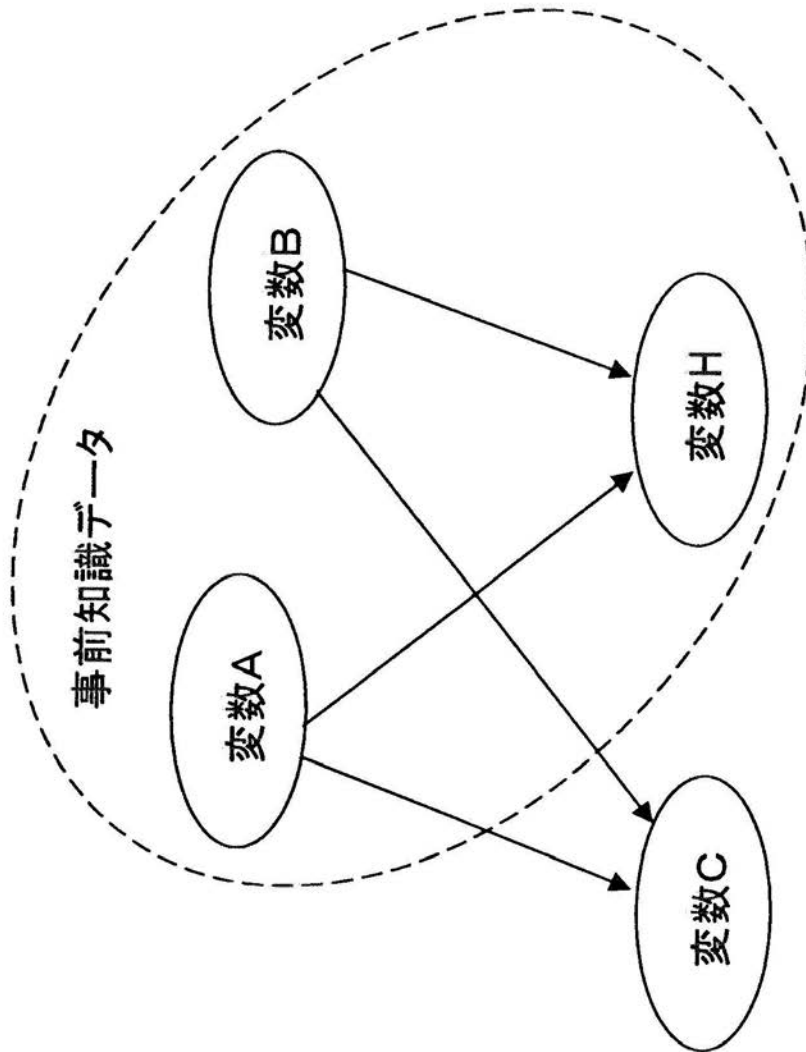
【図 4】

調査対象者ID	年齢	性別	選択基準
U0030	27	男	ブランド
U0056	29	男	値段
U0012	32	女	品質
・	・	・	・
・	・	・	・
・	・	・	・

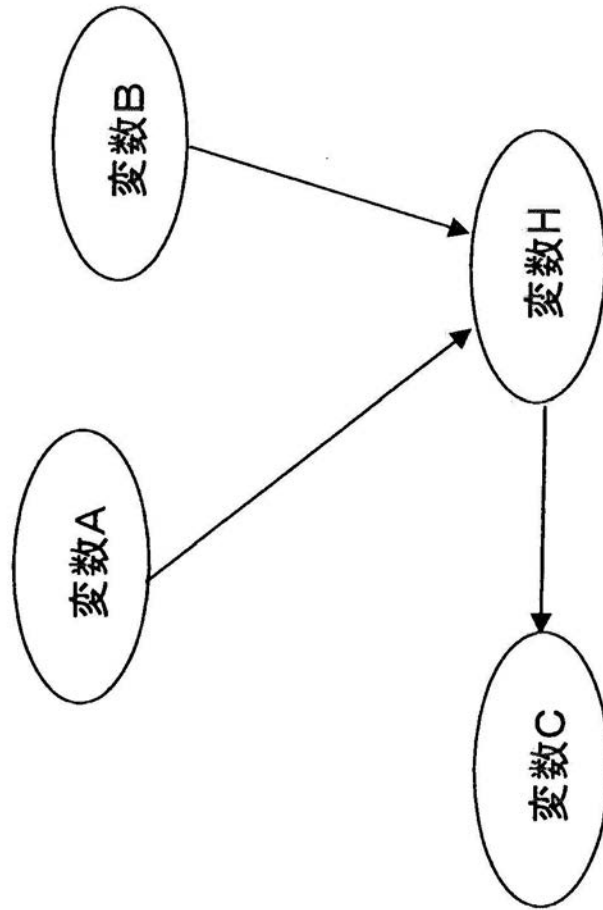
【図5】



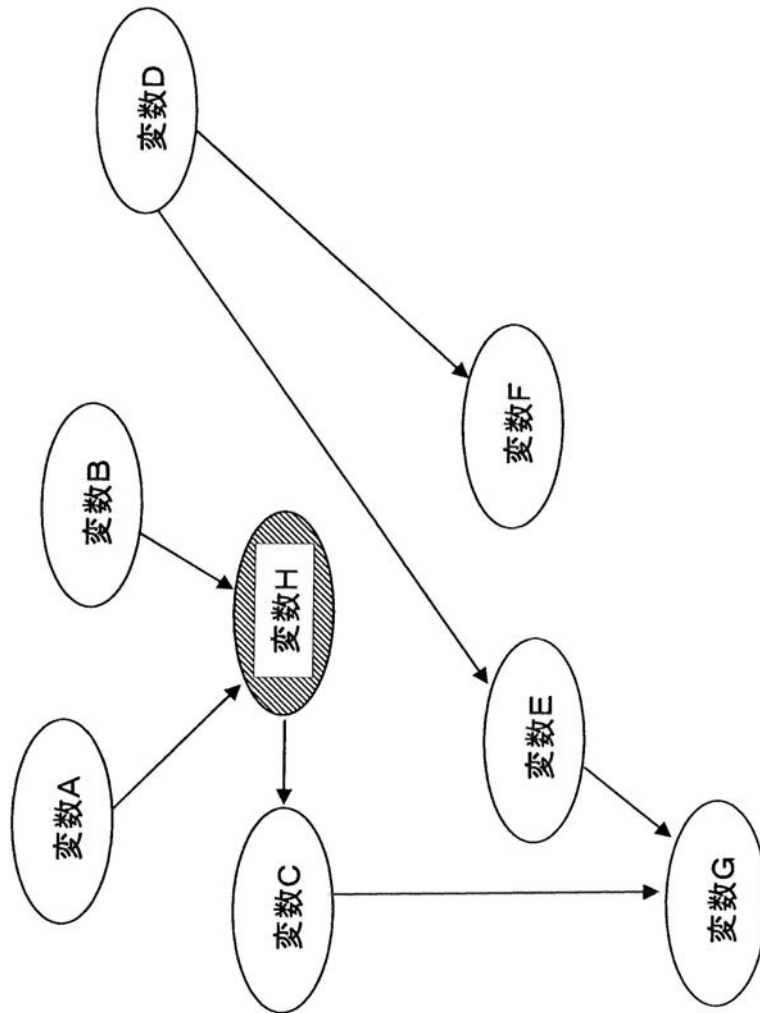
【図 6】



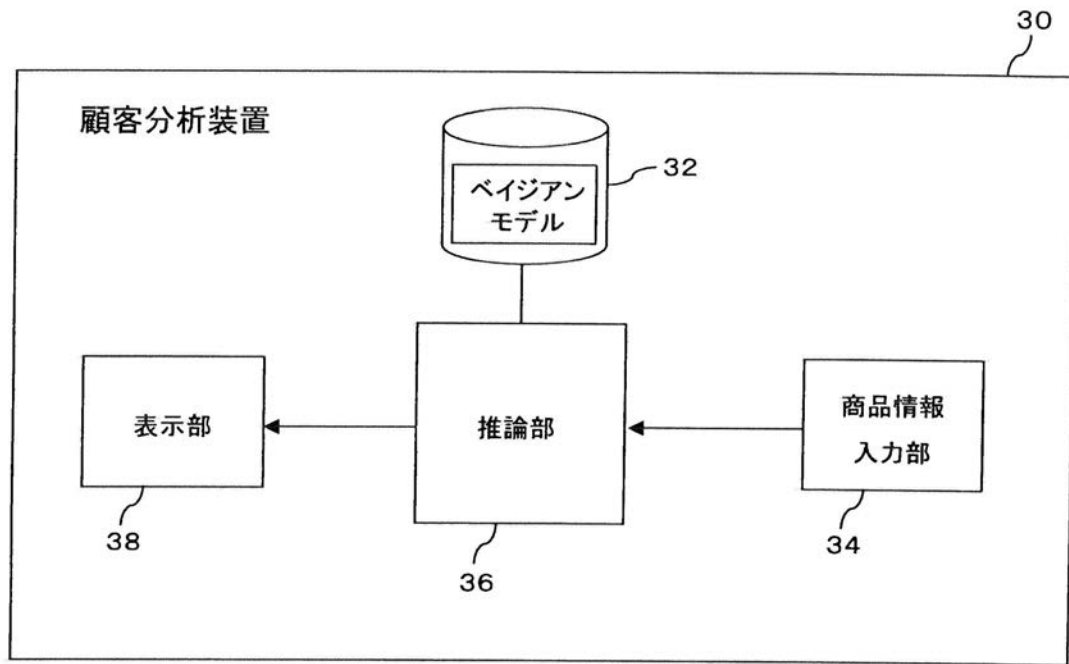
【図7】



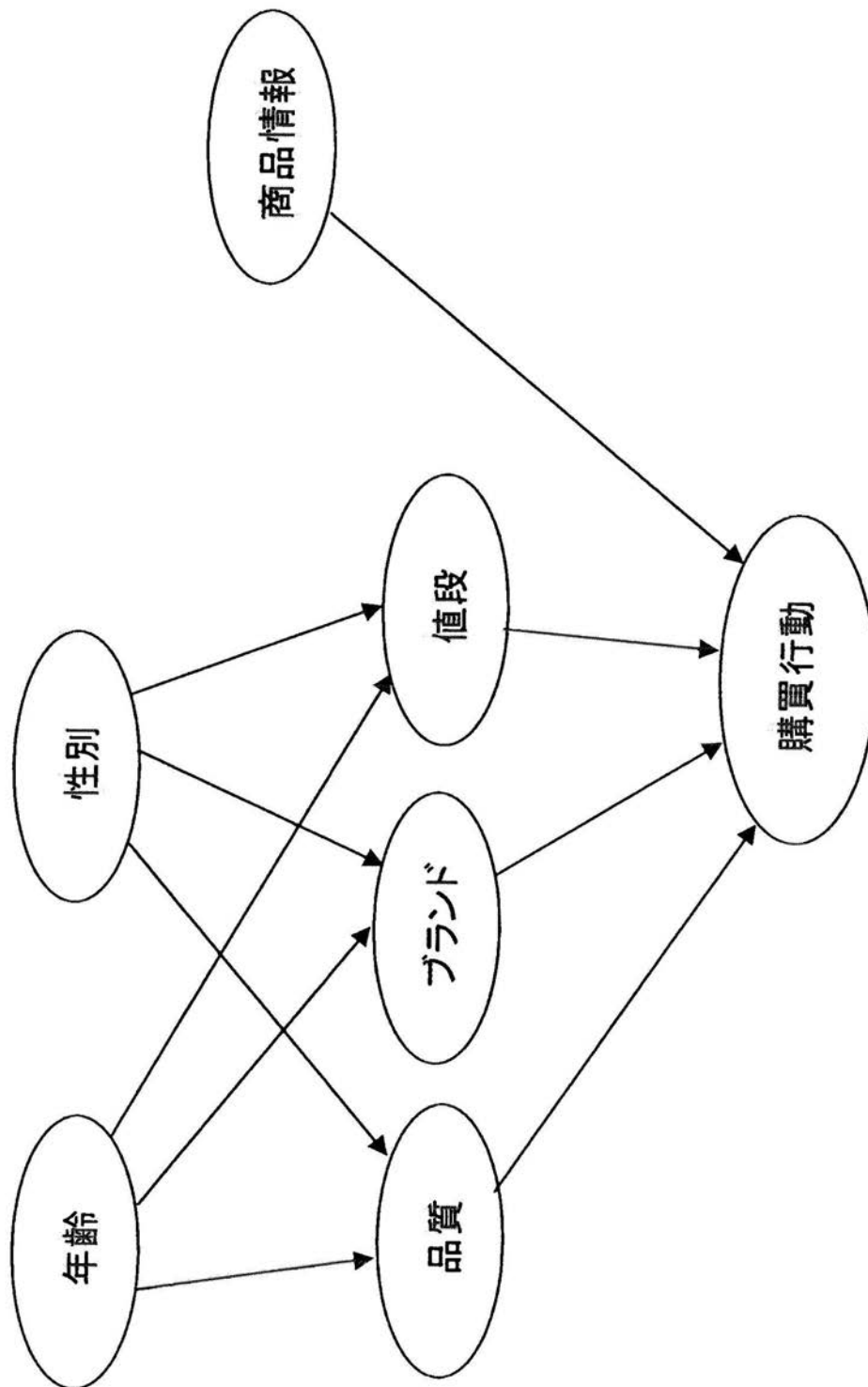
【 図 8 】



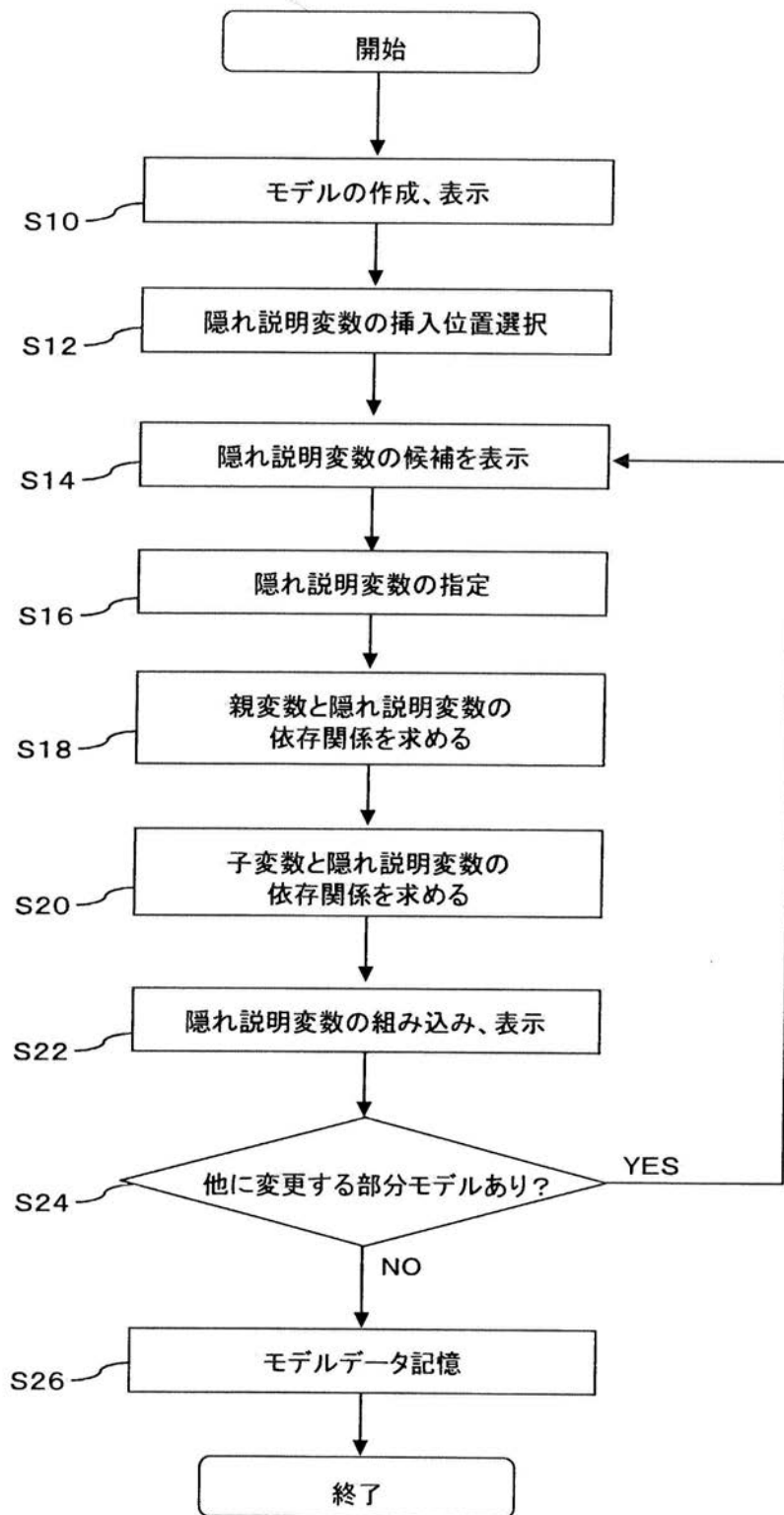
【図 9】



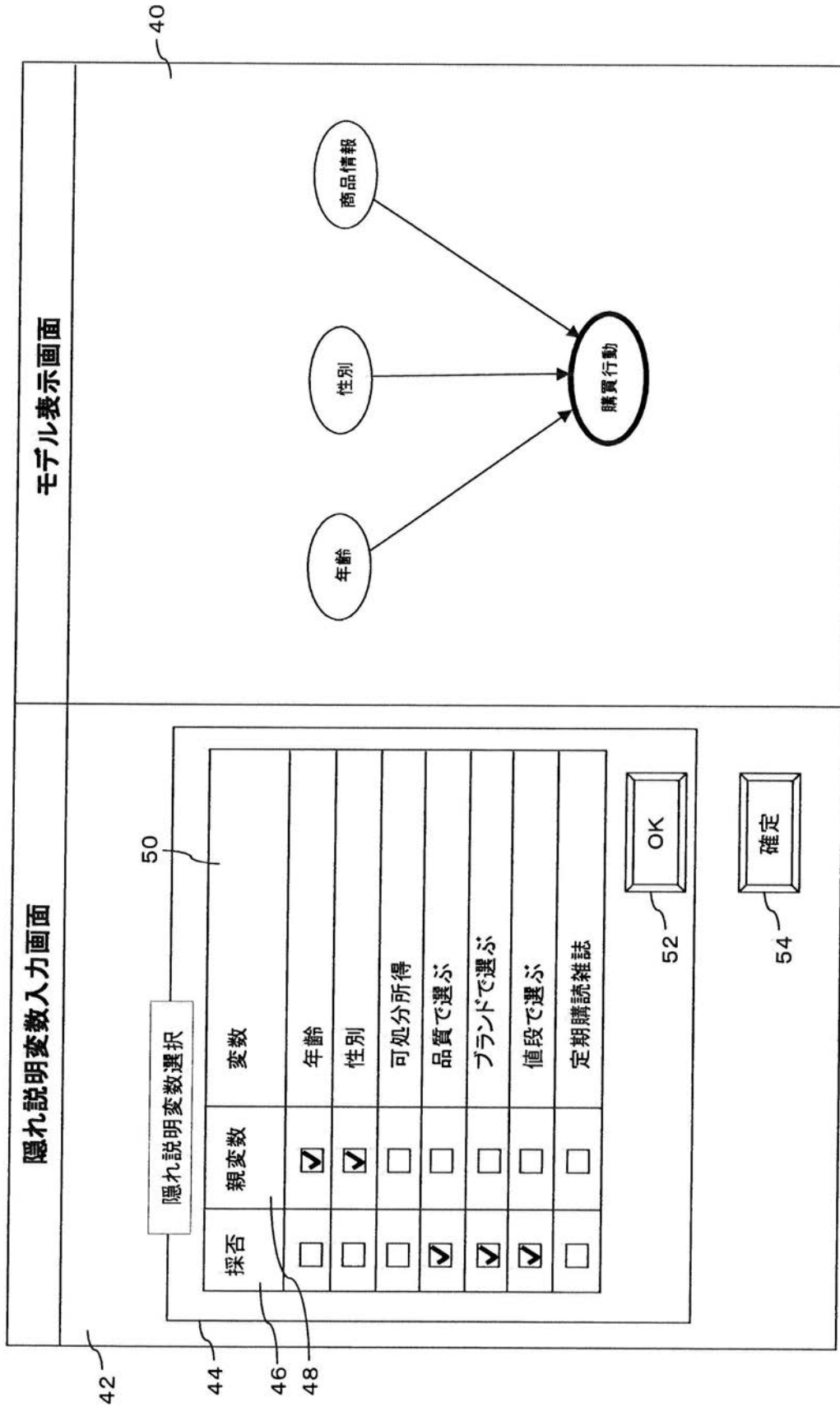
【図 10】



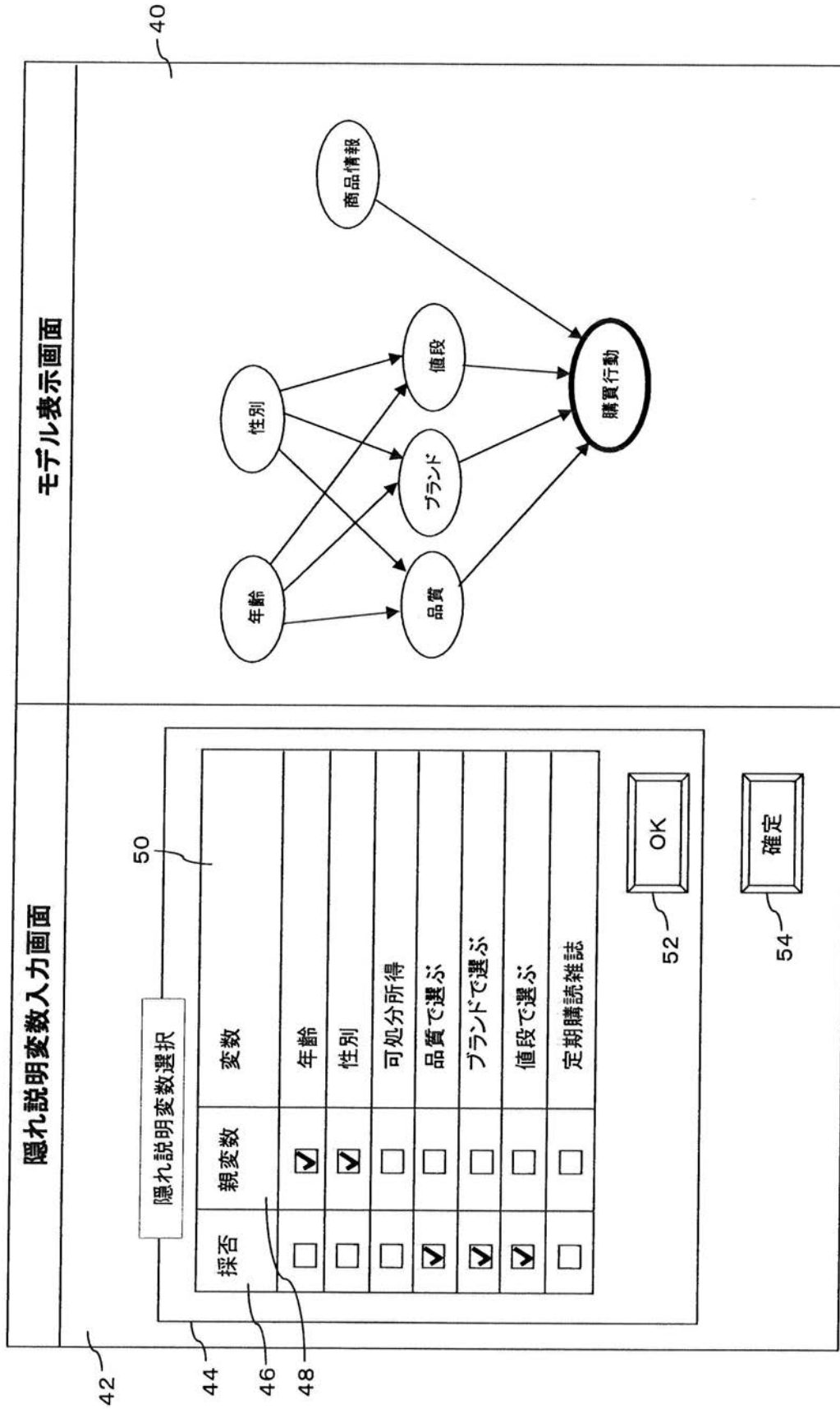
【図 11】



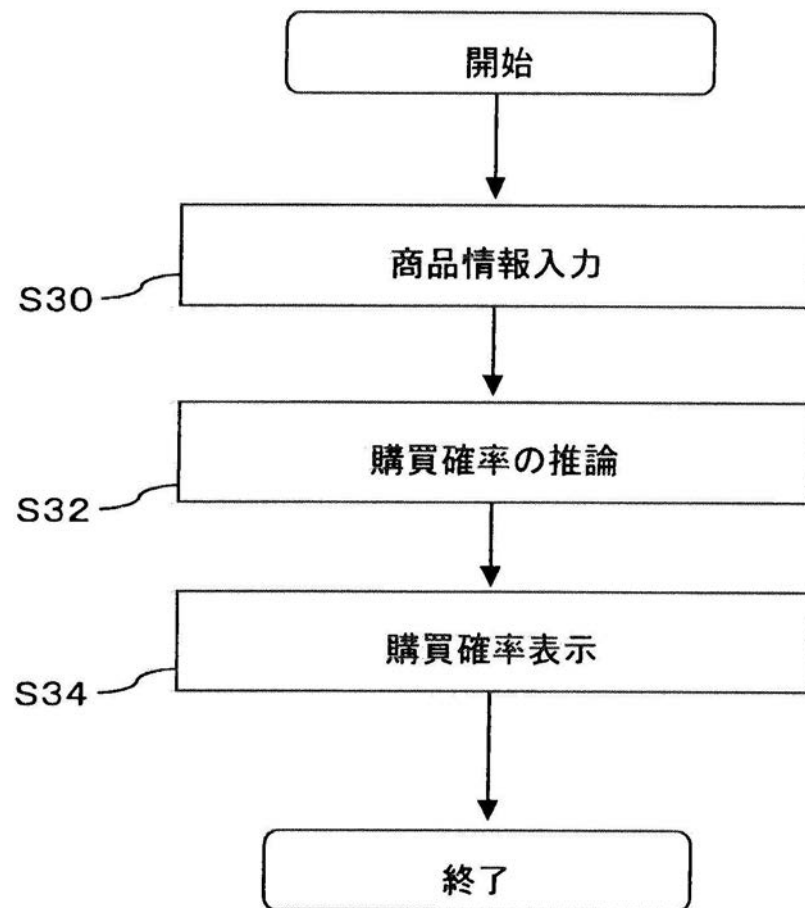
【図12】



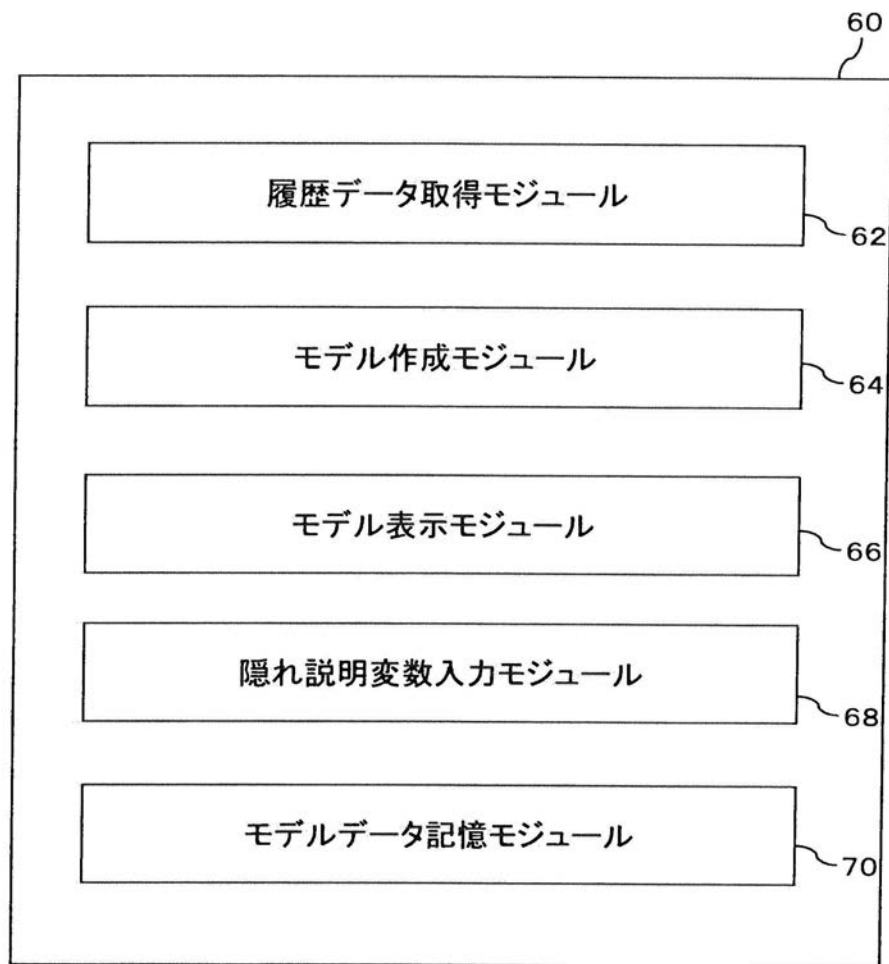
【図13】



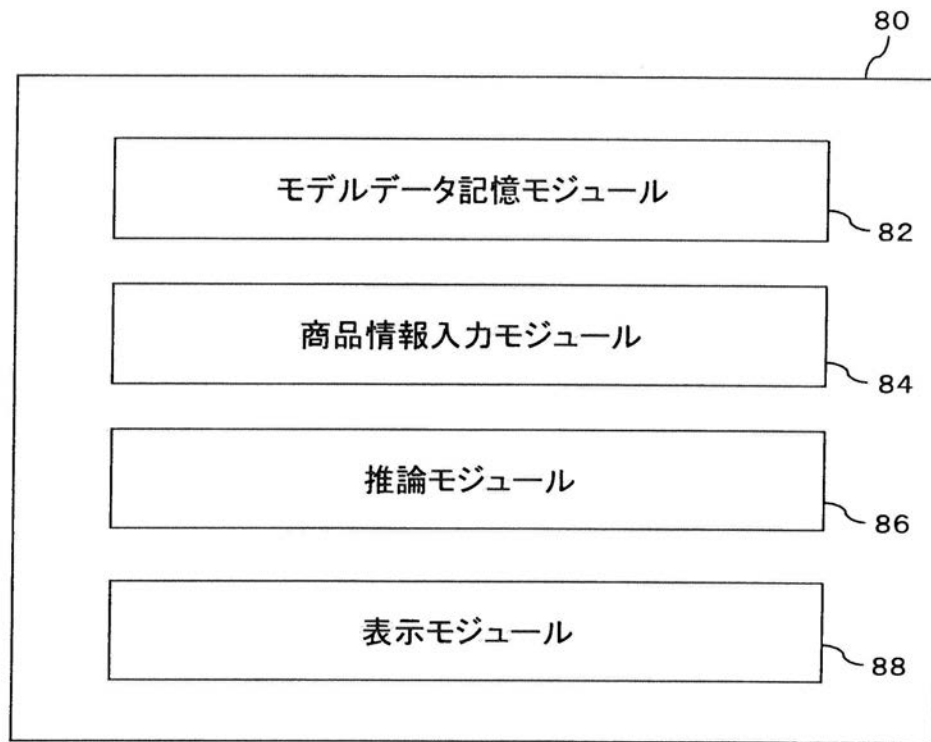
【図 14】



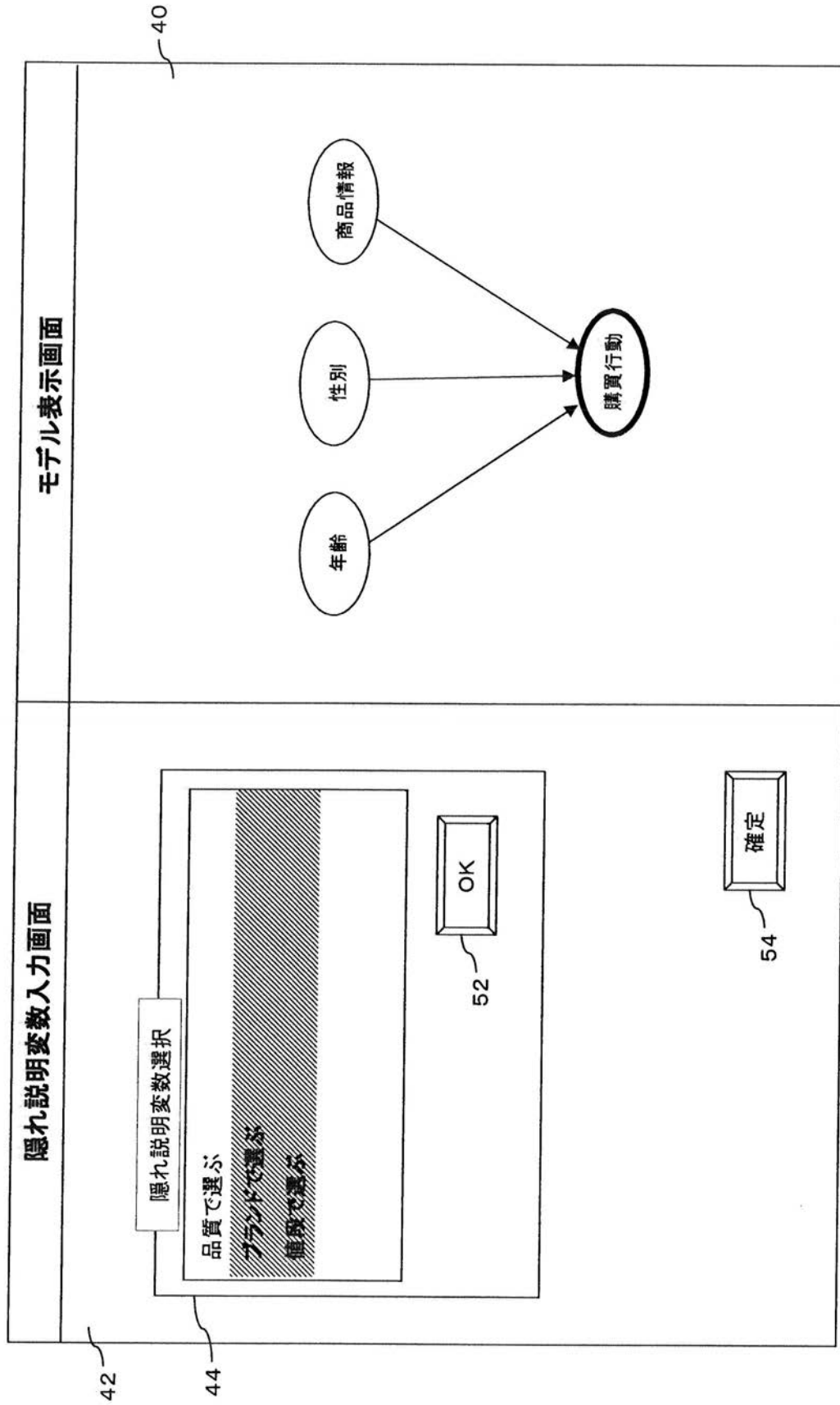
【図 15】



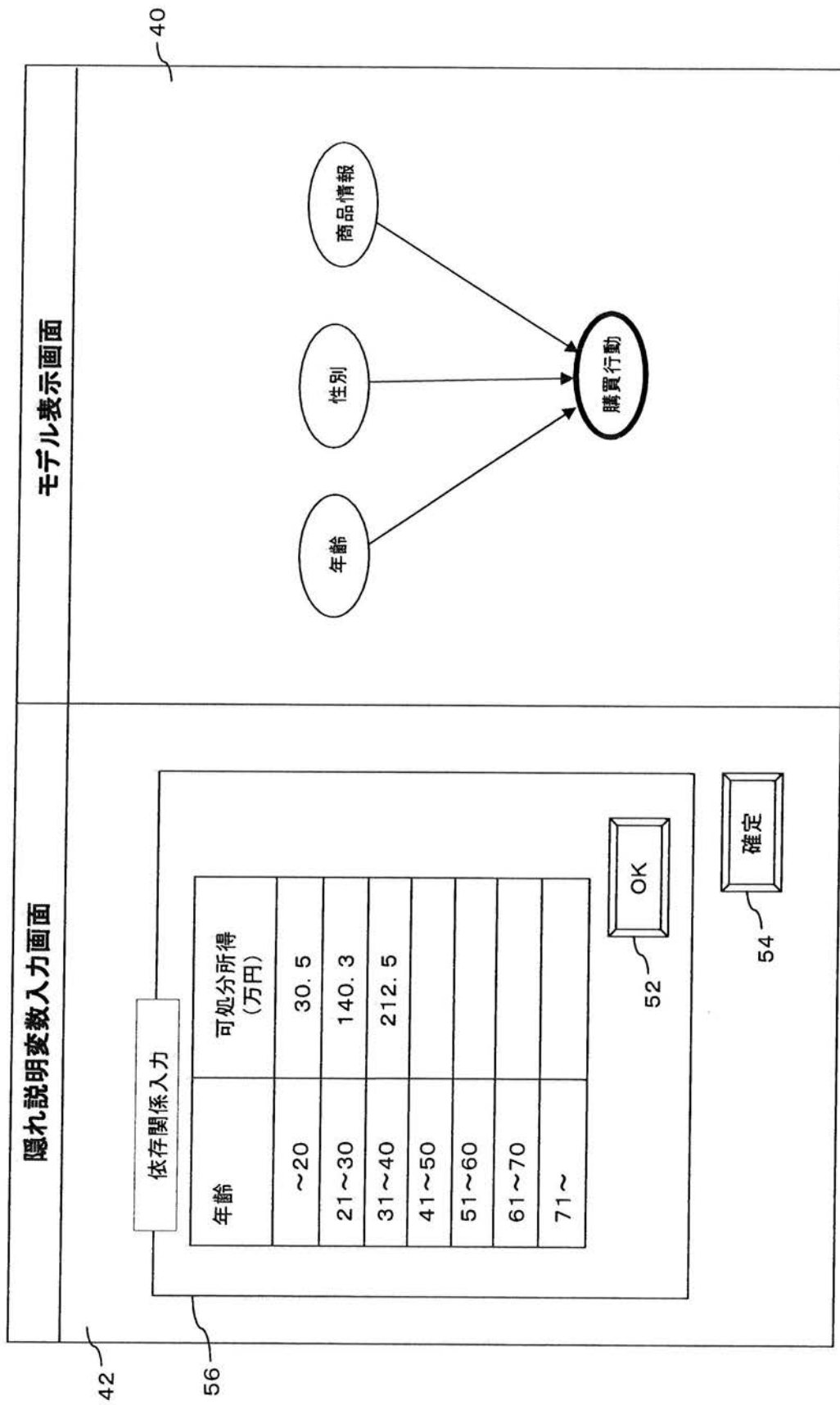
【図 16】



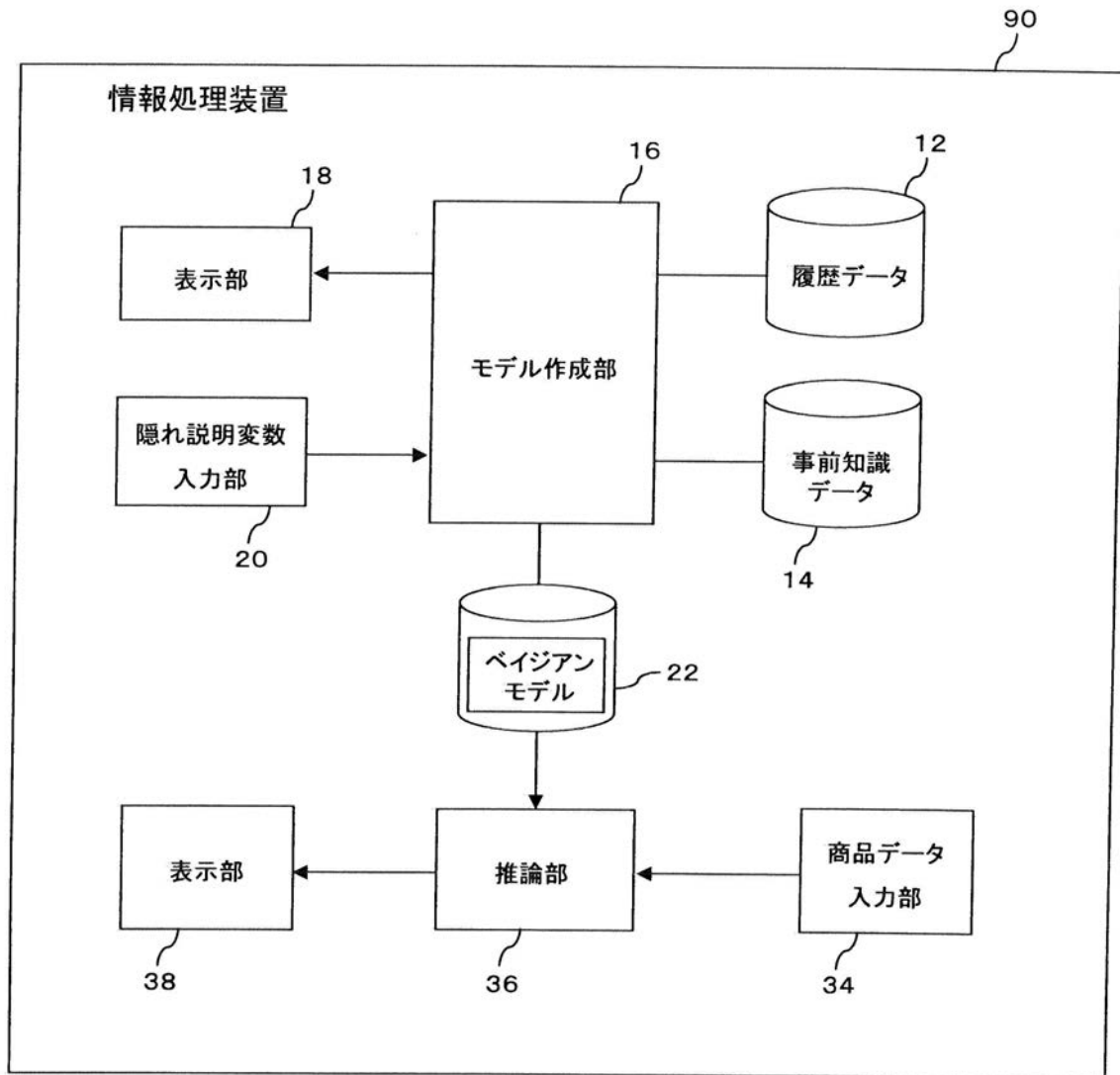
【図 17】



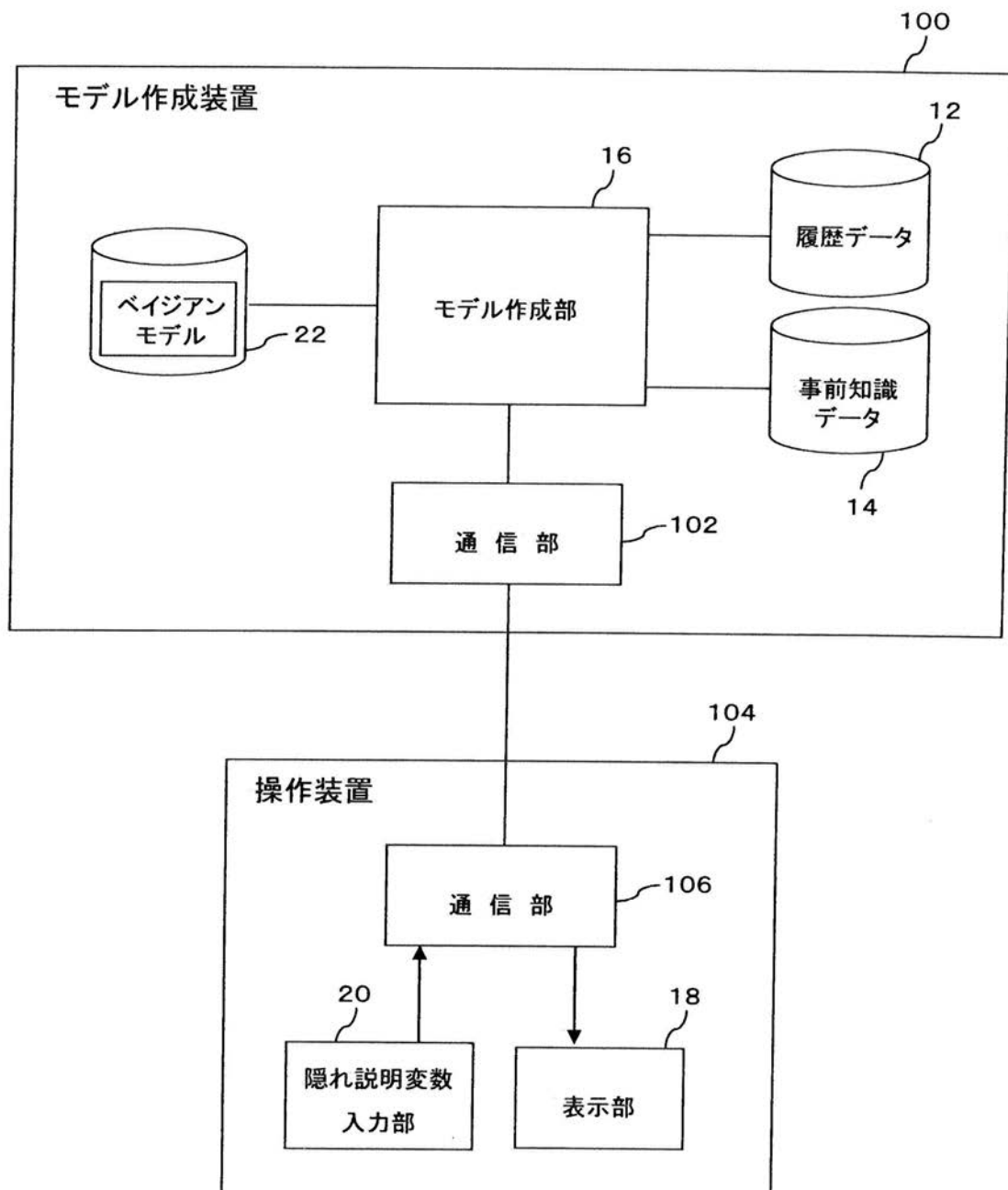
【図18】



【図 19】



【図 20】



フロントページの続き

特許法第30条第1項適用 平成15年8月10日 社団法人 計測自動制御学会発行の「計測と制御 第42巻 第8号」に発表

(74)代理人 100113549

弁理士 鈴木 守

(72)発明者 本村 陽一

東京都江東区青海2-41-6 独立行政法人産業技術総合研究所臨海副都心センター内

審査官 長谷川 篤男

(56)参考文献 Wai Lam, Bayesian network refinement via machine learning approach, IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998年 3月, Volume: 20, Issue: 3, pp.240-251

本村陽一 他, 確率ネットワークによるユーザモデル構築システム, IPA平成12年度成果論文集, 日本, 2001年11月14日

(58)調査した分野(Int.Cl., DB名)

G06N 5/04

JSTPlus(JDreamII)