

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2014-235654  
(P2014-235654A)

(43) 公開日 平成26年12月15日(2014.12.15)

(51) Int.Cl.

G06Q 10/06 (2012.01)

F I

G06Q 10/06 120

テーマコード (参考)

審査請求 未請求 請求項の数 6 O L (全 28 頁)

(21) 出願番号 特願2013-118091 (P2013-118091)  
(22) 出願日 平成25年6月4日(2013.6.4)

(71) 出願人 000003078  
株式会社東芝  
東京都港区芝浦一丁目1番1号  
(74) 代理人 110001612  
きさらぎ国際特許業務法人  
(72) 発明者 森 俊樹  
東京都港区芝浦一丁目1番1号 株式会社東芝内  
(72) 発明者 覚井 真吾  
東京都港区芝浦一丁目1番1号 株式会社東芝内  
(72) 発明者 田村 朱麗  
東京都港区芝浦一丁目1番1号 株式会社東芝内

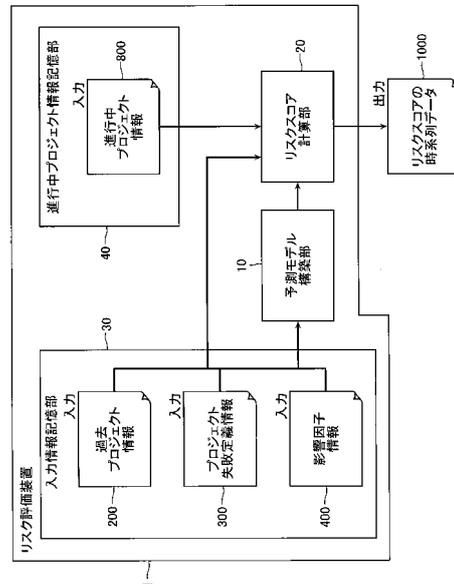
(54) 【発明の名称】 リスク評価装置

(57) 【要約】

【課題】 進行中プロジェクトにおけるプロジェクト失敗の危険度を定量化し、その時系列変化を求めて、プロジェクト失敗の兆候を早期に検知する。

【解決手段】 影響因子情報の区分情報に基づいて、進行中プロジェクト情報の各影響因子の値が属する区分に対応する区分値に置換することにより、符号化進行中プロジェクト情報を生成し、符号化過去プロジェクト情報と進行中プロジェクト情報に基づいて、プロジェクトの成功失敗に対する、影響因子の影響の度合いを示す情報である区分情報リフト値を各影響因子のタイミング情報及び区分値毎に算出し、符号化過去プロジェクト情報若しくは進行中プロジェクト情報に対して、各影響因子のタイミング情報、及び区分情報毎の区分情報リフト値を用いて、それぞれのプロジェクトについてタイミング情報の値ごとのリスクスコアを計算する。

【選択図】 図1



## 【特許請求の範囲】

## 【請求項 1】

過去に実施されたプロジェクトについて複数の影響因子の値を記述した情報である過去プロジェクト情報と、

過去に実施されたプロジェクトを、成功したプロジェクトと失敗したプロジェクトに区別するための基準を記述した情報であるプロジェクト失敗定義情報と、

影響因子の値が取り得る領域を複数の区分に分割した区分情報と、当該影響因子の値が取得可能な時期を示す情報であるタイミング情報とを有する情報である影響因子情報とを記憶する第 1 の記憶手段と、

プロジェクト失敗定義情報及び影響因子情報の区分情報に基づいて、過去プロジェクト情報の各影響因子の値が属する区分に対応する区分値に置換することにより、符号化過去プロジェクト情報を生成し、前記符号化過去プロジェクト情報に基づいて、プロジェクトの成功失敗に対する、影響因子の影響の度合いを示す情報である区分情報リフト値を各影響因子の区分値ごとに算出する第 1 の処理手段と、

進行中のプロジェクトについて複数の影響因子の値を記述した情報である進行中プロジェクト情報を記憶する第 2 の記憶手段と、

前記影響因子情報の区分情報に基づいて、前記進行中プロジェクト情報の各影響因子の値が属する区分に対応する区分値に置換することにより、符号化進行中プロジェクト情報を生成し、前記符号化過去プロジェクト情報と進行中プロジェクト情報に基づいて、プロジェクトの成功失敗に対する、影響因子の影響の度合いを示す情報である区分情報リフト値を各影響因子のタイミング情報及び区分値毎に算出し、符号化過去プロジェクト情報若しくは進行中プロジェクト情報に対して、各影響因子のタイミング情報、及び区分情報毎の区分情報リフト値を用いて、それぞれのプロジェクトについてタイミング情報の値ごとのリスクスコアを計算する第 2 の処理手段と

を有するリスク評価装置。

## 【請求項 2】

前記第 2 の処理手段は、影響変数が  $X_i$  であって、区分値が  $k$  である区分情報リフト値  $L_{X_i,k}$  を以下の式により算出し、

## 【数 2 3】

$$EL_{X_i} = \frac{3 \times 1.8182 + 6 \times 1.2571}{9} = 1.4441$$

(但し、 $c_1$ は、影響因子  $X_i$  について、区分値が  $k$  であって且つプロジェクトの成功失敗を示す格納フィールドの値が成功を示す値であるプロジェクトレコードの個数であり、

$c_2$ は、影響因子  $X_i$  について区分値が  $k$  以外且つ「NA」以外であって、且つプロジェクトの成功失敗を示す格納フィールドの値が成功を示す値であるプロジェクトレコードの個数であり、

$c_3$ は、影響因子  $X_i$  について区分値が  $k$  であって、且つプロジェクトの成功失敗を示す格納フィールドの値が失敗を示す値であるプロジェクトレコードの個数であり、

$c_4$ は、影響因子  $X_i$  について区分値が  $k$  以外且つ NA 以外であって、且つプロジェクトの成功失敗を示す格納フィールドの値が失敗を示す値であるプロジェクトレコードの個数であり、

$a$ 、 $b$ はそれぞれ任意の数である)

前記第 2 の処理手段は、以下の式により、各影響因子のタイミング情報、及び区分値毎の区分情報リフト値を用いてリスクスコアを計算する、

請求項 1 に記載のリスク評価装置

## 【数 2 4】

$$S_{p_0} = \text{Log}_{10} \left( \frac{\text{過去プロジェクト情報における成功を示す区分値を有するプロジェクト数}}{\text{過去プロジェクト情報の全プロジェクト数}} \right)$$

$$S_{p_j} = S_{p_{(j-1)}} + \sum_{T(X_i)=j} \text{Log}_{10}(L_{Xi_k})$$

(但し、 $S_{p_0}$ は、プロジェクトpのタイミング情報の値が0である場合のリスクスコア、 $S_{p_j}$ はプロジェクトpのタイミング情報の値がjである場合のリスクスコア、 $T(X_i)$ は影響因子 $X_i$ のタイミング情報の値、 $L_{Xi_k}$ は影響因子が $X_i$ で且つ区分値をkとする区分情報リフト値である)。

10

## 【請求項 3】

符号化過去プロジェクト情報のプロジェクトの成功失敗を示す値である成否判定値及びタイミング情報の値ごとの過去プロジェクトのリスクスコアに基づいて、プロジェクト失敗確率の回帰式を求め、求めたプロジェクト失敗確率回帰式を用いて、進行中プロジェクトに対するプロジェクト失敗確率を計算する第3の処理部をさらに有する、請求項1又は2に記載のリスク評価装置。

## 【請求項 4】

前記第3の処理部は、符号化された過去プロジェクト情報の成否判定値、及び、タイミング情報  $T = j$  における過去プロジェクトのリスクスコアを入力としてロジスティック回帰分析を実施し、下記式の  $a$ 、 $b$  を求めて失敗確率回帰式を得る、請求項3に記載のリスク評価装置

20

## 【数 2 5】

$$P_j = \frac{1}{1 + e^{-(a+bX)}}$$

(但し、 $P_j$ はタイミング情報  $T = j$  における、プロジェクト失敗の確率)。

## 【請求項 5】

影響因子及びその区分情報の候補の集合である影響因子候補情報を記憶する第3の記憶手段と、

30

前記影響因子候補情報を用いて過去プロジェクト情報から符号化過去プロジェクト情報を生成し、前記区分情報リフト値、又は区分情報リフト値が1以下の場合には前記区分情報リフト値の逆数に対して期待値をとって合計した値である期待区分情報リフト値を計算し、前記期待区分情報リフト値に基づいて前記影響因子情報に含める影響因子及びその区分情報を選択し、選択した影響因子及びその区分情報を含む影響因子情報を生成する第4の処理手段と

をさらに有する、請求項1又は2に記載のリスク評価装置。

## 【請求項 6】

前記第4の処理手段は、下記式により期待区分情報リフト値を計算する、請求項5に記載のリスク評価装置。

40

## 【数 2 6】

$$EL_{-Xi} = \frac{1}{N} \sum_k n_k |L_{Xi,k}|$$

但し、

$$N = \sum_k n_k \quad n_k : \text{区分情報 } k \text{ の数}$$

$$|L_{Xi,k}| = \begin{cases} L_{Xi,k} & (L_{Xi,k} \geq 1) \\ \frac{1}{L_{Xi,k}} & (L_{Xi,k} < 1) \end{cases}$$

10

但し、 $L_{Xi,k}$  : 影響因子  $X_i$ 、区分値  $k$  の影響区分情報リフト値

## 【発明の詳細な説明】

## 【技術分野】

## 【0001】

本明細書に記載の実施の形態は、リスク評価装置に関する。

## 【背景技術】

## 【0002】

20

大規模化、複雑化する開発プロジェクトにおいて、問題の予兆を早期にとらえ、プロジェクトの失敗を未然に防止するために、様々な開発管理データ（進捗、コスト、品質情報、など）の活用が重要となる。また、構成管理ツール、不具合管理ツールなど、様々な開発支援ツールの普及により、プロジェクトに関連する開発管理データを、開発担当者に余計な負担を掛けずに、収集、蓄積することが可能になってきた。

## 【0003】

プロジェクトの潜在的なリスクは目に見えにくく、対応が後手に回りがちであり、リスクを先回りした予見的な管理が必要である。変化が激しく、不確実性の高い状況においては、納期遅延、コスト超過、品質不足など、プロジェクトの失敗に関わるリスクを定量的に評価し、リスク予測値の変化を継続的にモニタリングしたいという要請があった。

30

## 【0004】

リスク予測の手法として以下のものが提案されていた。

## 【0005】

## (1) 重回帰分析による予測モデル

重回帰分析は、ソフトウェア開発管理における予測モデル構築において、最も一般的に用いられている手法の一つである。説明変数（入力）と目的変数（出力）の間の関係式を、直接、重回帰分析で求めて、説明変数の値が得られたときの目的変数の値そのものを予測する手法である。

## 【0006】

## (2) アンケート調査による混乱予測システム

過去プロジェクトに対して実施したアンケート調査の結果をベイズ識別器で学習し、新規プロジェクトに対する同アンケート結果に基づいて、そのプロジェクトが最終的に混乱状態に陥るかどうかを判定する手法である。

40

## 【先行技術文献】

## 【特許文献】

## 【0007】

【特許文献1】特開2011-076411号公報

【特許文献2】特開2010-108404号公報

【特許文献3】特開2007-265141号公報

【特許文献4】特開2009-187288号公報

50

【特許文献5】特開2011-141674号公報

【特許文献6】特開2008-204313号公報

【発明の概要】

【発明が解決しようとする課題】

【0008】

上記のような従来のリスク予見手法は以下のような問題点があった。

重回帰分析によるリスク予見手法については、外れ値や異常値の影響を受けやすいという問題がある。精度の高い予測モデルを構築するには、高精度かつ大量のデータが必要となるが、欠損値があると、サンプル数が減ってしまう。一般に、ソフトウェアの開発管理データは、ハードウェア製造プロセスの計測データ等に比べて、データの精度が低く、重回帰分析による予測モデル構築は困難な場合が多い。

10

【0009】

アンケート調査によるプロジェクト混乱予測システムにおいては、予測の実施はアンケート集計後の一時点のみであって、リスク予測値の更新は考慮されておらず、プロジェクト成功/失敗の判定のみを行う。また、失敗確率の計算は行わず、定量的評価を提供するものではない。また、離散値の変数のみを入力とし、連続値の変数やその閾値の決定方法には触れられていない。

【0010】

本発明の一の実施の形態は、進行中プロジェクトにおけるプロジェクト失敗の危険度を定量化し、その時系列変化を求めて、プロジェクト失敗の兆候を早期に検知する技術を提供することを目的とする。

20

【0011】

本発明の別の実施の形態は、進行中プロジェクトに対する、任意の時点でのプロジェクト失敗の確率を計算する技術を提供することを目的とする。

【0012】

本発明のさらに別の実施の形態は、最適な影響因子の選定、及び、区分情報の設定を行うことにより、予測モデル構築作業の手間を大幅に低減すると共に、プロジェクト失敗の予測精度向上を図る技術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】

【0013】

本発明の第1の実施の形態はリスク評価装置として提案される。このリスク評価装置は第1の記憶手段と、第1の処理手段と、第2の記憶手段と、第2の処理手段とを有する。

30

第1の記憶手段は、過去に実施されたプロジェクトについて複数の影響因子の値を記述した情報である過去プロジェクト情報と、過去に実施されたプロジェクトを、成功したプロジェクトと失敗したプロジェクトに区別するための基準を記述した情報であるプロジェクト失敗定義情報と、影響因子の値が取り得る領域を複数の区分に分割した区分情報と、当該影響因子の値が取得可能な時期を示す情報であるタイミング情報とを有する情報である影響因子情報とを記憶する。

【0014】

第1の処理手段は、プロジェクト失敗定義情報及び影響因子情報の区分情報に基づいて、過去プロジェクト情報の各影響因子の値が属する区分に対応する区分値に置換することにより、符号化過去プロジェクト情報を生成し、符号化過去プロジェクト情報に基づいて、プロジェクトの成功失敗に対する、影響因子の影響の度合いを示す情報である区分情報リフト値を各影響因子の区分値ごとに算出する。

40

【0015】

第2の記憶手段は、進行中のプロジェクトについて複数の影響因子の値を記述した情報である進行中プロジェクト情報を記憶する。

第2の処理手段は、影響因子情報の区分情報に基づいて、進行中プロジェクト情報の各影響因子の値が属する区分に対応する区分値に置換することにより、符号化進行中プロジェクト情報を生成し、符号化過去プロジェクト情報と進行中プロジェクト情報に基づいて

50

、プロジェクトの成功失敗に対する、影響因子の影響の度合いを示す情報である区分情報リフト値を各影響因子のタイミング情報及び区分値毎に算出し、符号化過去プロジェクト情報若しくは進行中プロジェクト情報に対して、各影響因子のタイミング情報、及び区分情報毎の区分情報リフト値を用いて、それぞれのプロジェクトについてタイミング情報の値ごとのリスクスコアを計算する。

【0016】

本発明の第2の実施の形態は、リスク評価装置として提案される。このリスク評価装置は符号化過去プロジェクト情報のプロジェクトの成功失敗を示す値である成否判定値及びタイミング情報の値ごとの過去プロジェクトのリスクスコアに基づいて、プロジェクト失敗確率の回帰式を求め、求めたプロジェクト失敗確率回帰式を用いて、進行中プロジェクトに対するプロジェクト失敗確率を計算する第3の処理部をさらに有する。

10

【0017】

本発明の第3の実施の形態はリスク評価装置として提案される。このリスク評価装置は影響因子及びその区分情報の候補の集合である影響因子候補情報を記憶する第3の記憶手段と、影響因子候補情報を用いて過去プロジェクト情報から符号化過去プロジェクト情報を生成し、区分情報リフト値、又は区分情報リフト値が1以下の場合には区分情報リフト値の逆数に対して期待値をとって合計した値である期待区分情報リフト値を計算し、期待区分情報リフト値に基づいて影響因子情報に含める影響因子及びその区分情報を選択し、選択した影響因子及びその区分情報を含む影響因子情報を生成する第4の処理手段とをさらに有する。

20

【図面の簡単な説明】

【0018】

【図1】第1の実施の形態に係るリスク評価装置の構成例を示す機能ブロック図

【図2】過去プロジェクト情報のデータ構成例を示す図

【図3】影響因子情報のデータ構成例を示す図

【図4】符号化過去プロジェクト情報のデータ構成例を示す図

【図5】区分情報リフト値を計算するための要素の関係を示す図

【図6】符号化過去プロジェクト情報における  $c_1$  ,  $c_2$  ,  $c_3$  ,  $c_4$  それぞれの値を示す図

【図7】符号化過去プロジェクト情報における  $c_1$  ,  $c_2$  ,  $c_3$  ,  $c_4$  それぞれの値を示す別の図

30

【図8】進行中プロジェクト情報のデータ例を示す図

【図9】符号化進行中プロジェクト情報のデータ構成例を示す図

【図10】各タイミング情報 ( $T=1, 2, 3, 4$ ) におけるリスクスコアの時系列変化を可視化した折れ線グラフを示す図

【図11】進行中プロジェクト (New1 ~ New5) のリスクスコアの時系列変化を示すグラフを示す図

【図12】リスクスコア算出処理の例を示したフローチャート

【図13】第2の実施の形態に係るリスク評価装置の構成例を示す機能ブロック図

【図14】ロジスティック回帰分析によるプロジェクト失敗確率回帰式を導出するための元データを示す図

40

【図15】プロジェクト失敗確率算出処理の例を示したフローチャート

【図16】第3の実施の形態に係るリスク評価装置の構成例を示す機能ブロック図

【図17】第3の実施の形態に係るリスク評価装置の主たる動作の例を示したフローチャート

【図18】第1、第2、第3の実施の形態を組み合わせたリスク評価装置の構成例を示す機能ブロック図

【発明を実施するための形態】

【0019】

以下、図面を参照して本発明の実施の形態に係るリスク評価装置を説明する。

50

## [ 0 . 用語の定義 ]

本明細書で使用する用語の定義を述べる。

## ( 1 ) プロジェクト情報

「プロジェクト情報」とは、プロジェクト実行中、または、終了後に、収集・記録された開発管理データをいう。「プロジェクト情報」は変数名と値のペアの集合として表現される。

## 【 0 0 2 0 】

## ( 2 ) プロジェクト失敗

「プロジェクト失敗」とは、ある評価基準により決定される、プロジェクト結果が望ましくない状態となったことをいう。「プロジェクト失敗」の典型的な例は、品質不良、コスト超過、納期遅延、などである。

10

## 【 0 0 2 1 】

## ( 3 ) 影響因子

「影響因子」とは、プロジェクト失敗に影響を及ぼす開発管理データとその属性情報。属性情報として、区分情報とタイミング情報がある。

## 【 0 0 2 2 】

## ( 4 ) 区分情報

「区分情報」とは、影響因子の属性情報の一つであって、影響因子に対応する変数の値域をいくつかの領域に分割したもとして定義される。

## 【 0 0 2 3 】

20

## ( 5 ) タイミング情報

「タイミング情報」とは、影響因子の属性情報の一つであって、影響因子に対応する変数の値が取得可能な時期（工程、フェーズ、マイルストーン、など）を示し、時期の早いものから順に、タイミング情報  $T = 1, 2, \dots, m$  とラベル付けされる。

## 【 0 0 2 4 】

## ( 6 ) 区分情報リフト値

「区分情報リフト値」とは、プロジェクト失敗のときに影響因子の値がある区分に属する確率を、条件がないときに影響因子の値が当該区分に属する確率で割った値をいう。「区分情報リフト値」は、その影響因子により得られる情報が、プロジェクト失敗の予測にどの程度寄与するかを示す。区分情報リフト値が「1」ならば、その影響因子はプロジェクト失敗の予測に寄与しないことを意味する。

30

## 【 0 0 2 5 】

## ( 7 ) リスクスコア

「リスクスコア」とは、ベイズ確率に基づき、各事象の独立性を仮定したときのプロジェクト失敗の事後確率を算出し、その対数をとることで得られる値を言う。リスクスコアの値が大きくなると、プロジェクト失敗の確率が高まっていると考えられる。

## 【 0 0 2 6 】

## ( 8 ) 期待区分情報リフト値

「期待区分情報リフト値」とは、区分情報リフト値、または、区分情報リフト値の逆数（区分情報リフト値が1以下の場合）に対して期待値をとって合計した値をいう。影響因子の区分情報の決定などに用いられる。

40

## 【 0 0 2 7 】

## ( 9 ) 予測モデル

「予測モデル」とは、プロジェクトの事前、または、進行中に得られるデータを入力として、将来の結果（例えば、納期、コスト、品質、など）を予測する数学モデルをいう。本明細書においては、プロジェクト失敗リスクの予測モデルを指す。

## 【 0 0 2 8 】

## ( 1 0 ) ベイズ統計

「ベイズ統計」とは、ベイズの定理を基礎とし、主観確率を積極的に取り入れた確率・統計論の体系。推測統計学など、主観確率を認めない確率・統計論の体系（頻度主義と呼

50

ばれる)と対比される。

【0029】

(11)単純ベイズ分類器

「単純ベイズ分類器」とは、ベイズ統計の応用の一つ。事後確率の計算に事象の独立性を仮定し、確率的分類を行う。スパムメールのフィルタリング技術などに応用されている。

【0030】

(12)ロジスティック回帰分析

「ロジスティック回帰分析」とは、統計的回帰モデルの一種であり、線形回帰分析が量的変数を予測するのに対して、ロジスティック回帰分析は2値の質的変数を従属変数として、説明変数を用いてその発生確率を予測する。

10

【0031】

[1.第1の実施の形態]

本発明の第1の実施の形態を説明する。第1の実施の形態は、リスク評価装置として提案される。

【0032】

第1の実施の形態にかかるリスク評価装置は、多様な測定データから構成される過去プロジェクト情報に基づいて、進行中プロジェクトにおけるプロジェクト失敗の危険度を、「リスクスコア」として定量化し、その時系列変化を提供する装置である。第1の実施の形態にかかるリスク評価装置は、ベイズ統計を応用して、過去のプロジェクトデータからプロジェクト失敗の影響因子とその影響の度合い(区分情報リフト値)を求め、進行中プロジェクトにおいて新しく得られたデータに基づき、この進行中プロジェクトのプロジェクト失敗の危険度を示すリスクスコアを算出する。第1の実施の形態にかかるリスク評価装置は、進行中プロジェクトの進行に応じて、リスクスコアを更新することができる。

20

【0033】

リスク評価装置は、コンピュータ、ワークステーションなどの情報処理装置であって、この情報処理装置は、演算処理装置(CPU)、主メモリ(RAM)、読み出し専用メモリ(ROM)、入出力装置(I/O)、及び必要な場合にはハードディスク装置等の外部記憶装置を具備している装置である。

【0034】

[1.1.第1の実施の形態に係るリスク評価装置の構成例]

図1に第1の実施の形態に係るリスク評価装置の構成例を示す機能ブロック図を掲げる。なお、機能ブロック図中に示す構成要素は、リスク評価装置の機能を機能ごとにまとめてブロックとして捉えたものであり、リスク評価装置が各構成要素に対応する基板、装置、回路、部品などの物理的構成要素を備えていなければならないことを意味するわけではない。また、「接続されている」とは、データ、情報、命令などの送受信、受け取り、受け渡しなどが可能な状態になっていることをいい、互いに配線で連結されているような物理的な接続に限られる意味ではない。本明細書中の他の機能ブロック図の説明についても同様である。

30

【0035】

リスク評価装置1は、入力情報記憶部30と、この入力情報記憶部30に接続された予測モデル構築部10と、この予測モデル構築部10に接続されたリスクスコア計算部20と、このリスクスコア計算部20へ接続された進行中プロジェクト情報記憶部40とを有する。入力情報記憶部30は第1の記憶手段に相当し、予測モデル構築部10は第1の処理手段に相当し、進行中プロジェクト情報記憶部40は第2の記憶手段に相当する。

40

【0036】

[1.1.1.入力情報記憶部]

入力情報記憶部30は、予測モデル構築に必要なデータを記憶する機能を有する。入力情報記憶部30は、過去プロジェクト情報200と、プロジェクト失敗定義情報300と、影響因子情報400を記憶する。

50

## 【 0 0 3 7 】

## [ 1 . 1 . 1 . 1 . 過去プロジェクト情報 ]

過去プロジェクト情報 2 0 0 は、過去に実施された複数の個別プロジェクト情報から構成される情報である。個別プロジェクト情報は、当該プロジェクトの実行中または終了後に、収集・記録された様々な開発管理データであり、変数の変数名と値のペアの集合として表される。変数は、連続変数でも離散変数でもよい。変数の値は一部、欠損値が含まれていても構わない。

## 【 0 0 3 8 】

図 2 に過去プロジェクト情報 2 0 0 のデータ構成例を示す。過去プロジェクト情報 2 0 0 は、個別プロジェクト情報に相当するレコード 2 0 1 を複数有している。図 2 に示す例では、過去プロジェクトPJ1から過去プロジェクトPJ10までの10件のプロジェクトに対応する10件のレコード 2 0 1 を有している。

10

## 【 0 0 3 9 】

各レコード 2 0 1 は、変数の値を格納するフィールドを有している。図 2 に示す例では、変数として「仕様レビュー指摘密度」と、「設計レビュー指摘密度」と、「設計レビュー効率」と、「単体テストバグ密度」と、「結合テストバグ密度」と、「出荷後バグ密度」とが採用されている。各レコード 2 0 1 は、「仕様レビュー指摘密度」格納フィールド 2 1 1 と、「設計レビュー指摘密度」格納フィールド 2 1 2 と、「設計レビュー効率」格納フィールド 2 1 3 と、「単体テストバグ密度」格納フィールド 2 1 4 と、「結合テストバグ密度」格納フィールド 2 1 5 と、「出荷後バグ密度」格納フィールド 2 1 6 とを有している。

20

## 【 0 0 4 0 】

各格納フィールド 2 1 1 ~ 2 1 6 は、該当するプロジェクトにおける対応する変数の値を格納する。変数の値が得られていない場合には、「欠損値」であることを示す情報が格納される。

## 【 0 0 4 1 】

## [ 1 . 1 . 1 . 2 . プロジェクト失敗定義情報 ]

プロジェクト失敗定義情報 3 0 0 は、過去に実施されたプロジェクトの集合を、“成功”したプロジェクト群と“失敗”したプロジェクト群に分類（区別）するための基準を定義した情報である。

30

## 【 0 0 4 2 】

典型的には、プロジェクト失敗定義情報 3 0 0 は、品質不良、コスト超過、納期遅延などに関連する結果指標（出荷後バグ率、コスト予実比、遅延日数、など）に対して、定量的な基準（閾値）を設定する情報である。

## 【 0 0 4 3 】

図 2 に示した過去プロジェクト情報 2 0 0 の例に対応するプロジェクト失敗定義情報 3 0 0 は、出荷後バグ密度の値が 4 . 5 以上であれば「失敗」、それ以外の値は「成功」とする。また、「失敗」、「成功」はそれぞれ後述する区分値が定められており、それぞれ「1」、「0」であるものとする。

## 【 0 0 4 4 】

## [ 1 . 1 . 1 . 3 . 影響因子情報 ]

影響因子情報 4 0 0 は、プロジェクトの成功失敗に影響すると思われる変数である影響因子の集合に対して、当該影響因子の値が取り得る領域をいくつかの領域に分割した「区分情報」と、当該影響因子の値が取得可能な時期（工程、フェーズ、マイルストーン、など）を示す情報である「タイミング情報」とで構成される。タイミング情報は、時期が早いものから順に、 $T = 1, 2, \dots, m$  としてよい（値の重複可）。

40

## 【 0 0 4 5 】

図 3 に影響因子情報 4 0 0 のデータ構成例を示す。影響因子情報 4 0 0 は、影響因子ごとに一つのレコード 4 0 1 を有する。各レコード 4 0 1 は、その影響因子を特定する情報（変数名、変数番号、変数IDなど）を格納する、影響因子特定情報格納フィールド 4 1 1

50

と、その影響因子の区分情報を格納する区分情報格納フィールド412と、その影響因子のタイミング情報を格納するタイミング情報格納フィールド413とを有する。

【0046】

「区分情報」は、そのレコードに該当する影響因子が取り得る値を複数に区分してできる区分(値域)を示す情報である。例えば、図3に示した例では、影響因子「使用レビュー指摘密度」の区分情報として、第1の区分(値域)が「[MIN, 2.5)」であり、第2の区分(値域)が「[2.5, MAX)」であり、第3の区分は「欠損値」であることが記録されている。なお、「[MIN, 2.5)」は最小値以上且つ2.5未満の値域を意味しており、「[2.5, MAX)」は2.5以上且つ最大値以下の値域を意味している。

【0047】

また、上記区分はそれぞれ「区分値」を有しており、例えば、上記第1の区分の区分値=0、上記第2の区分の区分値=1、上記第3の区分は区分値=NAである。この区分値は後述する過去プロジェクト情報の符号化に用いられる。

【0048】

また、図3に示す例では、説明のため「タイミング情報」として「T=1(要求仕様化工程)」のように表記したが、実際には値のみタイミング情報格納フィールド413に格納されていれば足りる。

【0049】

[1.1.2.予測モデル構築部]

予測モデル構築部10は、入力情報記憶部30に記憶されている情報から、過去プロジェクト情報200を符号化して、符号化過去プロジェクト情報を生成し、この符号化過去プロジェクト情報に基づいて区分情報リフト値を算出する機能を有する。

【0050】

[1.1.2.1.符号化過去プロジェクト情報]

符号化過去プロジェクト情報の生成について説明する。予測モデル構築部10は、プロジェクト失敗定義情報300、及び影響因子情報400の区分情報に基づいて、過去プロジェクト情報200を符号化する。「符号化」とは、それぞれのレコードの格納フィールドに格納されている値を区分値に置換することをいう。本実施の形態では、格納フィールド211から格納フィールド216までの値を、当該値に対応する区分値に置換する。格納フィールド211から格納フィールド215までの値は、影響因子情報400の区分情報に基づいて区分値に置換される。格納フィールド216の値はプロジェクト失敗定義情報300に基づいて区分値に置換される。

【0051】

図4に、図3に示した過去プロジェクト情報200を符号化した結果得られる、符号化過去プロジェクト情報のデータ構成例を示す。符号化過去プロジェクト情報500は、過去プロジェクト情報200と同じく、10件の過去プロジェクトPJ1~PJ10に対応する10件のレコード501を有している。各レコード501は、「仕様レビュー指摘密度」格納フィールド511と、「設計レビュー指摘密度」格納フィールド512と、「設計レビュー効率」格納フィールド513と、「単体テストバグ密度」格納フィールド514と、「結合テストバグ密度」格納フィールド515と、「出荷後バグ密度」格納フィールド516とを有している。

【0052】

予測モデル構築部10は、格納フィールド511から515のそれぞれに、過去プロジェクト情報200の対応する格納フィールド211から215に格納されている値に対応する区分値を格納する。例えば、過去プロジェクト情報200の過去プロジェクトPJ1に対応するレコード201において、仕様レビュー指摘密度格納フィールド211に格納されている値は「0.5」である。この値「0.5」は、影響因子情報400における仕様レビュー指摘密度に対応するレコード401の区分情報412に記述された第1の区分(値域)「[MIN, 2.5)」に含まれる。この第1の区分(値域)「[MIN, 2.5)」の区分値は「0」であるので、予測モデル構築部10は符号化過去プロジェクト情報500の過

10

20

30

40

50

去プロジェクトPJ1に対応するレコード501において、仕様レビュー指摘密度格納フィールド511に区分値「0」を格納する。同様に他の格納フィールド512から515にも対応する区分値を格納する。

【0053】

予測モデル構築部10は、出荷後バグ密度格納フィールド516に、対応する区分値を格納する。例えば、過去プロジェクト情報200の過去プロジェクトPJ1に対応するレコード201において、出荷後バグ密度格納フィールド216に格納されている値は「1.2」である。この値「1.2」は、出荷後バグ密度の値が4.5以上であるというプロジェクト失敗定義情報300に記述された失敗の条件に該当しないので「成功」と判定し、成功に対応する区分値「0」を、符号化過去プロジェクト情報500の過去プロジェクトPJ1に対応するレコード501において、出荷後バグ密度格納フィールド516にこの成功に対応する区分値「0」を格納する。

10

【0054】

予測モデル構築部10は、全てのレコード501の全ての格納フィールド511～516について対応する区分値を決定し、その区分値を格納フィールド511～516に格納する。

【0055】

図4に示す区分値の格納が完了した状態の符号化過去プロジェクト情報500である。なお、図4に示す例では、各格納フィールド511～516のラベルは、対応する影響因子(変数)名「仕様レビュー指摘密度」、「設計レビュー指摘密度」、「設計レビュー効率」、「単体テストバグ密度」、「結合テストバグ密度」、「出荷後バグ密度」に代えて、X1、X2a、X2b、X3、X4、Yとした。Xの右側の数字は当該影響因子のタイミング情報の値である。また、「設計レビュー指摘密度」、「設計レビュー効率」はともにタイミング情報T=2であるので、区別のためにさらに添え字a,bを付した。プロジェクトの成功/失敗を示す「出荷後バグ密度」はYと表記した。なお、これらの表記は後述する計算過程を説明するためであって、符号化過去プロジェクト情報500に係るデータ構成に限定される趣旨ではない。また、図4に示す例において「NA」は欠損値に対応する区分値である。

20

【0056】

[1.1.2.2.区分情報リフト値]

予測モデル構築部10は、前述の符号化過去プロジェクト情報200に基づいて、プロジェクトの成功/失敗に対する、影響因子の影響の度合いを示す情報である区分情報リフト値を各影響因子の区分値ごとに算出する。

30

【0057】

区分情報リフト値を算出するために、予測モデル構築部10は影響因子 $X_i$  ( $i$ はタイミング情報の値)の区分値 $k$  ( $0 \leq k \leq n$ )に対して、符号化過去プロジェクト情報500を参照して、以下の $c_1$ 、 $c_2$ 、 $c_3$ 、 $c_4$ を求める。

【0058】

「 $c_1$ 」は、影響因子 $X_i$ について、該当格納フィールドに格納されている値(区分値)が「 $k$ 」であって、且つプロジェクトの成功/失敗を示す格納フィールド516(Y)の値が0であるレコード501の個数である。

40

【0059】

「 $c_2$ 」は、影響因子 $X_i$ について、該当格納フィールドに格納されている値が「 $k$ 」以外且つ「NA」以外であって、且つプロジェクトの成功/失敗を示す格納フィールド516(Y)の値が成功を示す値「0」であるレコード501の個数である。

【0060】

「 $c_3$ 」は、影響因子 $X_i$ について、該当格納フィールドに格納されている値が「 $k$ 」であって、且つプロジェクトの成功/失敗を示す格納フィールド516(Y)の値が失敗を示す値「1」であるレコード501の個数である。

【0061】

「 $c_4$ 」は、影響因子 $X_i$ について、該当格納フィールドに格納されている値が「 $k$ 」以

50

外且つ「NA」以外であって、且つプロジェクトの成功/失敗を示す格納フィールド516(Y)の値が1であるレコード501の個数である。

【0062】

図5に $X_i$ 、 $k$ 、 $c_1$ 、 $c_2$ 、 $c_3$ 、 $c_4$ との関係を示す表を掲げる。

【0063】

図4に示した符号化過去プロジェクト情報500の例では、予測モデル構築部10は影響因子 $X_1$ 、 $X_{2a}$ 、 $X_{2b}$ 、 $X_3$ 、 $X_4$ のそれぞれの区分値「0」と「1」のそれぞれについて $c_1$ 、 $c_2$ 、 $c_3$ 、 $c_4$ を求めることになる。

【0064】

次に予測モデル構築部10は以下で定義される区分情報リフト値「 $L_{Xi_k}$ 」を計算する。  
区分情報リフト値「 $L_{Xi_k}$ 」の算出式を以下に示す。

【0065】

【数1】

$$L_{Xi_k} = \frac{(c_3+a)/(c_3+c_4+b)}{(c_1+c_3+a)/(c_1+c_2+c_3+c_4+b)} \dots \text{(式1)}$$

【0066】

上記式1において、分子は、プロジェクト失敗のとき影響因子 $X_i$ の取る値が区分値 $k$ に属する確率であり、分母は何も条件なしで、影響因子 $X_i$ の取る値が区分値 $k$ に属する確率である。

【0067】

上記式1における $a$ 、 $b$ は、データ数が少ないときに極端な値が出力されることを避けるための補正パラメータであり、一般に、 $a=1$ 、 $b=2$ （ラプラス補正）などが用いられる。また、データ数が十分にあれば、 $a=b=0$ （補正なし）でもよい。

【0068】

各影響因子の独立性を仮定したとき、プロジェクト失敗の事後確率は、ベイズの定理により、事前確率と、 $a=b=0$ （補正なし）のときの区分情報リフト値との積に一致する。

【0069】

区分情報リフト値の計算例を示す。図4に示した符号化過去プロジェクト情報500における影響因子= $X_1$ 且つ区分値=0についての区分情報リフト値 $L_{X1_0}$ は以下のとおりである。

【0070】

図6に影響因子= $X_1$ 且つ区分値=0での、符号化過去プロジェクト情報500における $c_1$ 、 $c_2$ 、 $c_3$ 、 $c_4$ それぞれの値を示す。すなわち、 $c_1=4$ 、 $c_2=2$ 、 $c_3=1$ 、 $c_4=2$ となる。これを上記式1にあてはめる。

【0071】

【数2】

$$L_{X1_0} = \frac{(1+1)/(1+2+2)}{(4+1+1)/(4+2+1+2+2)} = 0.7333$$

【0072】

(但し、ラプラス補正あり： $a=1$ 、 $b=2$ )

すなわち、影響因子= $X_1$ 且つ区分値=0についての区分情報リフト値 $L_{X1_0}=0.7333$ を得る。

【0073】

区分情報リフト値の別の計算例を示す。図4に示した符号化過去プロジェクト情報500における影響因子= $X_1$ 且つ区分値=1についての区分情報リフト値 $L_{X1_1}$ は以下のと

おりである。

【0074】

図7に影響因子=X1且つ区分値=0での、符号化過去プロジェクト情報500におけるc1, c2, c3, c4それぞれの値を示す。すなわち、c1=2, c2=4, c3=2, c4=1となる。これを上記式1にあてはめる。

【0075】

【数3】

$$L\_X1\_1 = \frac{(2+1)/(2+1+2)}{(2+2+1)/(2+4+2+1+2)} = 1.32$$

10

【0076】

(但し、ラプラス補正あり: a = 1, b = 2)

すなわち、影響因子=X1且つ区分値=1についての区分情報リフト値L\_X1\_1=1.32を得る。

【0077】

予測モデル構築部10は、他の影響因子X2a、X2b、X3、X4のそれぞれの区分値「0」と「1」についても同様に区分情報リフト値を算出し、各区分情報リフト値を出力する。

【0078】

以上で予測モデル構築部10の説明を終了する。

20

【0079】

[1.1.3.進行中プロジェクト情報記憶部]

進行中プロジェクト情報記憶部40は、進行中プロジェクト情報800を記憶する機能を有する。進行中プロジェクト情報800は、1つ以上の“進行中”(成功/失敗が未確定)のプロジェクト情報の集まりである。進行中プロジェクト情報800は、前述の過去プロジェクト情報200と同様のデータ構成を有している。

【0080】

図8に進行中プロジェクト情報のデータ例を示す。進行中プロジェクト情報800は、個別の進行中プロジェクト情報に相当するレコード801を複数有している。図8に示す例では、進行中プロジェクトNEW1から進行中プロジェクトNEW5までの5件の進行中プロジェクトに対応する5件のレコード801を有している。

30

【0081】

各レコード801は、過去プロジェクト情報200と同様に、変数の値を格納するフィールドを有している。図8に示す例では、変数として「仕様レビュー指摘密度」と、「設計レビュー指摘密度」と、「設計レビュー効率」と、「単体テストバグ密度」と、「結合テストバグ密度」と、「出荷後バグ密度」とが採用されている。各レコード801は、「仕様レビュー指摘密度」格納フィールド811と、「設計レビュー指摘密度」格納フィールド812と、「設計レビュー効率」格納フィールド813と、「単体テストバグ密度」格納フィールド814と、「結合テストバグ密度」格納フィールド815と、「出荷後バグ密度」格納フィールド816とを有している。

40

【0082】

各格納フィールド811~816は、該当するプロジェクトにおける対応する変数の値を格納する。変数の値が得られていない場合には、「欠損値」であることを示す情報が格納される。また、該当する変数についての検証・検査等が見実施である場合には「未実施」であることを示す情報が格納される。

【0083】

なお、進行中プロジェクトの“プロジェクト完了後”(成功/失敗が確定した後)のプロジェクト情報は、過去プロジェクト情報200の中に含めて、将来の進行中プロジェクトの予測に活用することも可能である。

【0084】

50

## [ 1.1.4. リスクスコア計算部 ]

リスクスコア計算部 20 は、過去及び進行中プロジェクトに対するリスクスコアの計算し、過去及び進行中プロジェクトに対するリスクスコアの時系列データを出力する機能を有する。リスクスコア計算部 20 は、影響因子情報 400 と符号化過去プロジェクト情報 500 と、各影響因子の区分情報毎の区分情報リフト値と、進行中プロジェクト情報とを入力情報として用いる。

## 【 0085 】

## [ 1.1.4.1. 進行中プロジェクト情報の符号化 ]

リスクスコア計算部 20 は、進行プロジェクト情報 800 を符号化して、符号化進行中プロジェクト情報 800 を生成する。符号化処理の内容は、予測モデル構築部 10 による過去プロジェクト情報の符号化と同様であって、影響因子情報 400 の区分情報に基づいて、進行中プロジェクト情報 800 の各格納フィールド 811 ~ 815 に格納されている値を、対応する区分値に置換する。ただし、進行中のプロジェクトにおいては、プロジェクトの成功 / 失敗は、未確定なので、プロジェクト失敗定義情報 300 はここでは必要ない。

10

## 【 0086 】

図 9 に符号化進行中プロジェクト情報 900 のデータ構成例を示す。図 9 に示す例は、図 8 に示した進行中プロジェクト情報 800 から生成された符号化進行中プロジェクト情報 900 の例である。

## 【 0087 】

符号化進行中プロジェクト情報 900 は、進行中プロジェクト情報 800 と同じく、5 件の進行中プロジェクト NEW1 ~ NEW5 に対応する 5 件のレコード 901 を有している。各レコード 901 は、「仕様レビュー指摘密度」格納フィールド 911 と、「設計レビュー指摘密度」格納フィールド 912 と、「設計レビュー効率」格納フィールド 913 と、「単体テストバグ密度」格納フィールド 914 と、「結合テストバグ密度」格納フィールド 915 と、「出荷後バグ密度」格納フィールド 916 とを有している。

20

## 【 0088 】

リスクスコア計算部 20 は、格納フィールド 911 から 915 のそれぞれに、進行中プロジェクト情報 800 の対応する格納フィールド 811 から 815 に格納されている値に対応する区分値を格納する。但し、進行中プロジェクト情報 800 で格納されている値が「未実施」を示す値である場合は、符号化進行中プロジェクト情報 900 の対応する格納フィールドにも「未実施」を示す値を格納する。

30

## 【 0089 】

なお、図 9 に示す例では、各格納フィールド 911 ~ 916 のラベルは、対応する影響因子 (変数) 名「仕様レビュー指摘密度」、「設計レビュー指摘密度」、「設計レビュー効率」、「単体テストバグ密度」、「結合テストバグ密度」、「出荷後バグ密度」に代えて、X1、X2a、X2b、X3、X4、Yとしている。

## 【 0090 】

## [ 1.1.4.2. リスクスコアの計算 ]

リスクスコア計算部 20 は、符号化過去プロジェクト情報 900 と進行中プロジェクト情報 800 に基づいて、各影響因子のタイミング情報、及び区分値毎の区分情報リフト値を用いてリスクスコアを計算する。

40

## 【 0091 】

リスクスコア計算部 20 は、プロジェクト p (過去、進行中を問わない) のリスクスコアを以下のように計算する。

(1) プロジェクト p の初期リスクスコア  $S_{p\_0}$  を以下の式 2 より算出する。

## 【 0092 】

## 【数 4】

$$S_{p_0} = \text{Log}_{10} \left( \frac{\text{過去プロジェクト情報における} Y=1 \text{であるプロジェクト数}}{\text{過去プロジェクト情報の全プロジェクト数}} \right) \dots (\text{式2})$$

## 【0093】

一般的には、当該プロジェクトに関して特に初期情報がないと仮定して、上記のように初期リスクスコア  $S_{p_0}$  を設定するが、もし、プロジェクト毎に初期リスクスコアを変えたい場合には、個別に設定してもよい。

(2) プロジェクト  $p$  のタイミング情報  $T=j$  についてのリスクスコアを算出する。

10

## 【0094】

プロジェクト  $p$  の影響因子  $X_i$  の値を  $k$  とする区分情報リフト値を  $L_{Xi_k}$ 、影響因子  $X_i$  のタイミング情報の値を  $T(X_i)$  とすると、リスクスコア計算部 20 は、プロジェクト  $p$  の  $T=j$  におけるリスクスコア  $S_{p_j}$  を以下の式 3 により計算する。

## 【0095】

## 【数 5】

$$S_{p_j} = S_{p_{(j-1)}} + \sum_{T(X_i)=j} \text{Log}_{10}(L_{Xi_k}) \dots (\text{式3})$$

20

## 【0096】

但し、 $k = NA$  ならば、 $L_{Xi_k} = 1$  とする。式 3 の右辺第 2 項は、タイミング情報  $T=j$  である全ての影響因子についての区分情報リフト値の対数の合計を意味する。また、式 3 における  $k$  の値は、符号化過去プロジェクト情報 500 又は符号化進行中プロジェクト情報 900 のプロジェクト  $p$  に対応するレコード 501 又はレコード 901 の影響変数  $X_i$  に対応する格納フィールドに格納された区分値である。

## 【0097】

進行中プロジェクトが過去プロジェクトと同様の傾向を示すと仮定すれば、一般に、プロジェクトのリスクスコアが高いほど、プロジェクト失敗の確率が高いと判断できる。

## 【0098】

30

リスクスコア計算の例を挙げる。図 8 及び図 9 に示した進行中プロジェクト情報 800 及び符号化進行中プロジェクト情報 900 に含まれるプロジェクト New1 のリスクスコアの計算例を示す。

## 【0099】

タイミング情報  $T = 1$  についてのプロジェクト New1 のリスクスコア  $S_{New1_1}$  は以下のように算出される。

## 【0100】

## 【数 6】

$$\begin{aligned} S_{New1_1} &= S_{New1_{(1-1)}} + \sum_{T(X_i)=1} \text{Log}_{10}(L_{Xi_k}) \\ &= S_{New1_0} + \text{Log}_{10}(L_{X1_0}) \\ &= \text{Log}_{10}(0.3) + \text{Log}_{10}(0.7333) \\ &= -0.6578 \end{aligned}$$

40

## 【0101】

なお、上記式中の区分情報リフト値  $L_{X1_0}$  の値は、前述の予測モデル構築部 10 によって算出された値を使用する。

また、タイミング情報  $T = 2$  についてのプロジェクト New1 のリスクスコア  $S_{New1_2}$  は以下のように算出される。

50

【 0 1 0 2 】

【 数 7 】

$$\begin{aligned}
 S\_New1\_2 &= S\_New1\_(2-1) + \sum_{T(Xi)=2} \text{Log}_{10}(L\_Xi\_k) \\
 &= S\_New1\_1 + \text{Log}_{10}(L\_X2a\_1) + \text{Log}_{10}(L\_X2b\_NA) \\
 &= -0.6578 + \text{Log}_{10}(1.44) + \text{Log}_{10}(1) \\
 &= -0.4992
 \end{aligned}$$

【 0 1 0 3 】

10

なお、本例においてタイミング情報T=2である影響因子はX2a,X2bの2つであるため、この2つの影響因子の区分情報リフト値の対数を計算に含めている。上記式中の区分情報リフト値L\_X2a\_1の値は、前述の予測モデル構築部10によって算出された値を使用する。区分情報リフト値L\_X2a\_NAはk=NAの場合の値=1を用いる。

【 0 1 0 4 】

タイミング情報T = 3についてのプロジェクトNew1のリスクスコアS\_New1\_3は以下のよう  
に算出される。

【 0 1 0 5 】

【 数 8 】

$$\begin{aligned}
 S\_New1\_3 &= S\_New1\_(3-1) + \sum_{T(Xi)=3} \text{Log}_{10}(L\_Xi\_k) \\
 &= S\_New1\_2 + \text{Log}_{10}(L\_X3\_1) \\
 &= -0.4992 + \text{Log}_{10}(1.6667) \\
 &= -0.2774
 \end{aligned}$$

20

なお、上記式中の区分情報リフト値L\_X3\_1の値は、前述の予測モデル構築部10によっ  
て算出された値を使用する。

【 0 1 0 6 】

タイミング情報T = 4についてのプロジェクトNew1のリスクスコアS\_New1\_4は以下のよ  
うに算出される。

30

【 0 1 0 7 】

【 数 9 】

$$\begin{aligned}
 S\_New1\_4 &= S\_New1\_(4-1) + \sum_{T(Xi)=4} \text{Log}_{10}(L\_Xi\_k) \\
 &= S\_New1\_3 + \text{Log}_{10}(L\_X4\_0) \\
 &= -0.2774 + \text{Log}_{10}(0.7333) \\
 &= -0.4121
 \end{aligned}$$

40

【 0 1 0 8 】

なお、上記式中の区分情報リフト値L\_X4\_0の値は、前述の予測モデル構築部10によっ  
て算出された値を使用する。

【 0 1 0 9 】

以上で進行プロジェクトNew1について、タイミング情報T=1、T=2、T=3、T=4までのそれ  
ぞれについてリスクスコアを求めることができた。これで各タイミング情報(T=1,...,4)  
におけるリスクスコアの時系列変化を示すデータを生成、出力することが可能となる。

【 0 1 1 0 】

図10に各タイミング情報(T=1, 2, 3, 4)におけるリスクスコアの時系列変化を可視  
化した折れ線グラフを示す。このグラフより、T=2(設計工程)からT=3(単体テスト工程

50

）にかけて、プロジェクト失敗リスクの予測値が少しずつ上昇していることが分かる。

【 0 1 1 1 】

同様に、全ての進行中プロジェクト（New1～New5）についてリスクスコアを算出することができる。各進行中プロジェクトのリスクスコアをタイミング情報（ $T=1, 2, 3, 4$ ）毎に示すことで、各進行中プロジェクトのリスクスコアの時系列変化をグラフで可視化することができる。図 1 1 に進行中プロジェクト（New1～New5）のリスクスコアの時系列変化を示すグラフを掲げる。このグラフから、例えば、以下のような情報が分かる。プロジェクト New3、及び、New4 は、プロジェクト失敗リスクの予測値であるリスクスコアが上昇傾向にある。これらプロジェクトについては監視強化が望ましく、必要に応じて、これらプロジェクトについてリスク軽減策を実施すべき、などの判断が行える。

10

【 0 1 1 2 】

また、プロジェクト New5 は、 $T=2$ （設計工程）時点では、比較的风险が低いとみなせる。

【 0 1 1 3 】

[ 1 . 2 . リスク評価装置の動作例 ]

図 1 に示したリスク評価装置 1 の動作例を説明する。図 1 2 にリスク評価装置 1 の主たる動作であるリスクスコア算出処理の例を示したフローチャートを示す。

【 0 1 1 4 】

リスクスコア算出処理において、リスク評価装置 1、より詳しくは予測モデル構築部 1 0 は過去プロジェクト情報 2 0 0、プロジェクト失敗定義情報 3 0 0、影響因子情報 4 0 0 から符号化過去プロジェクト情報 5 0 0 を生成する（S10）。

20

【 0 1 1 5 】

次にリスク評価装置 1、より詳しくは予測モデル構築部 1 0 は符号化過去プロジェクト情報 5 0 0 及び影響因子情報 4 0 0 に基づいて、各影響因子について区分値ごとに区分情報リフト値を算出する（S20）。

【 0 1 1 6 】

次にリスク評価装置 1、より詳しくはリスクスコア計算部 2 0 は、あらかじめ記憶されている進行プロジェクト情報 8 0 0、影響因子情報 4 0 0 に基づいて、符号化進行中プロジェクト情報 9 0 0 を生成する（S30）。

【 0 1 1 7 】

次にリスク評価装置 1、より詳しくはリスクスコア計算部 2 0 は、符号化過去プロジェクト情報 9 0 0 若しくは進行中プロジェクト情報 8 0 0 に対して、各影響因子のタイミング情報、及び区分情報毎の区分情報リフト値を用いて、それぞれのプロジェクトについてタイミング情報の値ごとのリスクスコアを計算する（S40）。

30

【 0 1 1 8 】

図 2 に示した過去プロジェクト情報 2 0 0 に基づいてリスク評価装置 1 が出力するリスクスコアを  $S_{p_j}$  で表すと、プロジェクト PJ1 について： $S_{PJ1_0}$ 、 $S_{PJ1_1}$ 、 $S_{PJ1_2}$ 、 $S_{PJ1_4}$  という値の組となる。同様に、他のプロジェクトのそれぞれについてもリスク評価装置 1 はタイミング情報の値ごとのリスクスコア  $S_{p_0}$ 、 $S_{p_1}$ 、 $S_{p_2}$ 、...、 $S_{p_m}$ （ $m$  はタイミング情報  $T$  がとる最大の値）を出力する。

40

【 0 1 1 9 】

リスク評価装置 1 は計算の結果得られたリスクスコアを数値の列挙の形式で出力しても良いし、グラフとして出力するようにしてもかまわない。出力形式に限定はない。

以上でリスク評価装置 1 の動作例の説明を終了する。

【 0 1 2 0 】

[ 1 . 3 . 第 1 の実施の形態の利点 ]

（1）多様の測定データから構成される過去プロジェクト情報に基づいて、進行中プロジェクトにおけるプロジェクト失敗の危険度を、複数の測定データを総合した「リスクスコア」として定量化し、その時系列変化を求めることができる。

【 0 1 2 1 】

50

(2) プロジェクト失敗の危険度を表す「リスクスコア」の時系列変化を監視することにより、プロジェクト失敗の兆候を早期に検知し、適切なタイミングで対応策を実施することができる。

【0122】

[2. 第2の実施の形態]

本発明の第2の実施の形態を説明する。第2の実施の形態はリスク評価装置として提案される。

【0123】

第2の実施の形態にかかるリスク評価装置は、第1の実施の形態の特徴に加えて、リスクスコアを入力とするロジスティック回帰分析により、任意の時点でのプロジェクト失敗の確率を計算することを特徴とする。

10

【0124】

第2の実施の形態にかかるリスク評価装置は、コンピュータ、ワークステーションなどの情報処理装置であって、この情報処理装置は、演算処理装置(CPU)、主メモリ(RAM)、読み出し専用メモリ(ROM)、入出力装置(I/O)、及び必要な場合にはハードディスク装置等の外部記憶装置を具備している装置である。

【0125】

[2.1. リスク評価装置の構成例]

図13に第2の実施の形態に係るリスク評価装置の構成例を示す機能ブロック図を掲げる。

20

【0126】

リスク評価装置2は、入力情報記憶部30と、この入力情報記憶部30に接続された予測モデル構築部10と、この予測モデル構築部10に接続されたリスクスコア計算部20と、このリスクスコア計算部20へ接続された進行中プロジェクト情報記憶部40と、リスクスコア計算部20に接続されたプロジェクト失敗確率計算部50を有する。プロジェクト失敗確率計算部50は第3の処理手段に相当する。

【0127】

リスク評価装置2は、第1の実施の形態に係るリスク評価装置1と基本的同様の構成要素を有しており、プロジェクト失敗確率計算部50をさらに有している点が異なっている。

30

【0128】

リスク評価装置2の構成要素のうち、第1の実施の形態に係るリスク評価装置1の構成要素と同一のものは、同一の参照符号を付してそれら構成要素の詳細な説明は省略する。

【0129】

[2.1.1. プロジェクト失敗確率計算部]

プロジェクト失敗確率計算部50は、リスクスコア計算部20が出力するリスクスコアの時系列データと、符号化過去プロジェクト情報500とに基づいて、過去プロジェクトに基づく失敗確率回帰式の導出を行い、この失敗確率回帰式を用いて進行中プロジェクトに対する失敗確率の計算を行い、タイミング情報毎( $T = 1, 2, \dots, m$ )のプロジェクト失敗確率を出力する機能を有する。

40

【0130】

なお、リスクスコアの時系列データについては、過去プロジェクト分、及び、進行中プロジェクト分の両方を用いる。また、符号化過去プロジェクト情報500に必要な情報は、 $Y$  (プロジェクトの成功 / 失敗) の値のみである。

【0131】

[2.1.1.1. 過去プロジェクトに基づく失敗確率回帰式の導出]

プロジェクト失敗確率計算部50による、過去プロジェクトに基づく失敗確率回帰式の導出について説明する。

【0132】

プロジェクト失敗確率計算部50は、符号化過去プロジェクト情報500のプロジェクト

50

トの成功 / 失敗を示す値を格納する格納フィールド（図5に示す例では格納フィールド 5 1 6）の値（以下「成否判定値」と呼ぶ）、及び、タイミング情報毎（ $T = 1, 2, \dots, m$ ）の過去プロジェクトのリスクスコアに基づいて、プロジェクト失敗確率の回帰式を求める。

【 0 1 3 3 】

具体的には、プロジェクト失敗確率計算部 5 0 は、符号化された過去プロジェクト情報 5 0 0 の成否判定値、及び、タイミング情報  $T = j$  における過去プロジェクトのリスクスコアを入力としてロジスティック回帰分析を実施し、下記式 4 の a, b を求める。なお、ロジスティック回帰分析は、MINITABやRなど、標準的な統計ツールに実装されているものを用いてよい。

【 0 1 3 4 】

【 数 1 0 】

$$p_j = \Pr(Y=1|X) = \frac{1}{1+e^{-(a+bX)}} \quad \dots(\text{式 4})$$

【 0 1 3 5 】

上記式 4 において、「 $P_j$ 」はタイミング情報  $T = j$  における、成否判定値（ $Y$ ） = 1（失敗）の確率を意味し、「 $X$ 」は過去プロジェクト Z のリスクスコア  $S\_Z\_j$  を意味する。

【 0 1 3 6 】

元データによっては、ロジスティック回帰分析による式の導出が行えない場合もある。そのときは、該当するタイミング情報での失敗確率は求められない。例えば、元データが完全分離可能な場合（ある閾値で成功 / 失敗が完全に識別できる）、などである。

【 0 1 3 7 】

他のタイミング情報（ $T = j$  以外）についても、同様の計算を実施する。

【 0 1 3 8 】

失敗確率回帰式の導出の例を示す。図 1 4 に、タイミング情報  $T=4$  の場合の、ロジスティック回帰分析によるプロジェクト失敗確率回帰式を導出するための元データを示す。この元データ 1 4 0 0 は、各過去プロジェクトの成否判定値（格納フィールド 5 1 6 に格納されている値）と、対応する過去プロジェクト p のリスクスコア  $S\_p\_4$  とで構成される。

【 0 1 3 9 】

プロジェクト失敗確率計算部 5 0 は、元データ 1 4 0 0 より、 $a=3.544$ 、 $b=7.342$  という結果を得え、以下の失敗確率回帰式を得る。

【 0 1 4 0 】

【 数 1 1 】

$$p_j = \Pr(Y=1|X) = \frac{1}{1+e^{-(3.554+7.342X)}} \quad \dots(\text{式 5})$$

【 0 1 4 1 】

上記式 5 によって得られるプロジェクト失敗確率は、タイミング情報  $T=4$  の場合に適用されるものとなる。プロジェクト失敗確率計算部 5 0 は、他のタイミング情報についてもそれぞれ失敗確率回帰式を得る。

【 0 1 4 2 】

[ 2 . 1 . 1 . 1 . 失敗確率回帰式によるプロジェクト失敗確率の計算 ]

プロジェクト失敗確率計算部 5 0 は、求めたプロジェクト失敗確率回帰式を用いて、進行中プロジェクトに対するプロジェクト失敗確率を計算する。求めたいタイミング情報の失敗確率回帰式を用いて、対象プロジェクトのリスクスコアを入力すれば、そのプロジェクトにおけるプロジェクト失敗確率が得られる。

【 0 1 4 3 】

プロジェクト失敗確率の算出例を示す。図 8 に示した例における進行中プロジェクト New1 について、タイミング情報  $T=4$  の場合のプロジェクト失敗確率  $P\_New1\_4$  を算出するため

10

20

30

40

50

には、上記式 5 を用いた以下の計算を行う。

【 0 1 4 4 】

【 数 1 2 】

$$P_{\_New1\_4} = \frac{1}{1 + e^{-(3.554 + 7.342 \times (-0.4121))}} = 0.6258$$

【 0 1 4 5 】

なお、上記式において $X = -0.4121$ （進行中プロジェクトNew1のタイミング情報 $T=4$ のリスクスコア；リスクスコア計算部 20 によって算出されている）を用いている。

10

【 0 1 4 6 】

同様にして、プロジェクト失敗確率計算部 50 は、いずれの進行中プロジェクトNew1, New2, New3, New4, New5についても、任意のタイミング情報（ $T=1, \dots, 4$ ）におけるプロジェクト失敗確率を計算可能である。

【 0 1 4 7 】

[ 2 . 2 . リスク評価装置 2 の動作例 ]

リスク評価装置 2 の動作例について説明する。図 15 に第 2 の実施の形態に係るリスク評価装置 2 の主たる動作であるプロジェクト失敗確率算出処理の例を示したフローチャートを示す。

【 0 1 4 8 】

20

プロジェクト失敗確率算出処理において、リスク評価装置 2 は、リスクスコア算出処理を実行する（S110）。リスクスコア算出処理は、第1の実施の形態におけるリスクスコア算出処理（図 12、ステップS10～S40参照）と同一の処理であるので、処理内容の詳述は省略する。

【 0 1 4 9 】

次にリスク評価装置 2、より詳しくはプロジェクト失敗確率計算部 50 は、過去プロジェクト情報 200 に基づいて失敗確率回帰式の導出を実行する（S120）。

【 0 1 5 0 】

次にリスク評価装置 2、より詳しくはプロジェクト失敗確率計算部 50 は、ステップS120において導出された失敗確率回帰式を用いて、進行中プロジェクトのプロジェクト失敗確率を算出する（S130）。

30

【 0 1 5 1 】

最後に、リスク評価装置 2、より詳しくはプロジェクト失敗確率計算部 50 は、ステップS130において算出したプロジェクト失敗確率を出力する。

以上でリスク評価装置 2 はプロジェクト失敗確率算出処理を終了する。

【 0 1 5 2 】

[ 2 . 3 . 第 2 の実施の形態の利点 ]

（1）進行中プロジェクトが過去プロジェクトと同様の傾向を示すと仮定することにより、ロジスティック回帰分析を用いて、進行中プロジェクトに対する、任意の時点でのプロジェクト失敗の確率を計算することができる。

40

【 0 1 5 3 】

（2）プロジェクト失敗の確率は、直観的に理解しやすく、プロジェクト制御のための意思決定等に有効に役立てることができる。

[ 3 . 第 3 の実施の形態 ]

本発明の第 3 の実施の形態を説明する。第 3 の実施の形態はリスク評価装置として提案される。

【 0 1 5 4 】

第 3 の実施の形態にかかるリスク評価装置は、第 1 の実施の形態の特徴に加えて、期待区分情報リフト値を用いて、影響因子の選定や区分情報の設定などを行い、影響因子情報を生成することを特徴とする。

50

## 【 0 1 5 5 】

第3の実施の形態にかかるリスク評価装置は、コンピュータ、ワークステーションなどの情報処理装置であって、この情報処理装置は、演算処理装置（CPU）、主メモリ（RAM）、読み出し専用メモリ（ROM）、入出力装置（I/O）、及び必要な場合にはハードディスク装置等の外部記憶装置を具備している装置である。

## 【 0 1 5 6 】

## [ 3 . 1 . リスク評価装置の構成例 ]

図16に第3の実施の形態に係るリスク評価装置の構成例を示す機能ブロック図を掲げる。リスク評価装置3は、影響因子候補情報記憶部60と、影響因子候補情報記憶部60に接続された影響因子情報決定部70と、影響因子情報決定部70に接続された入力情報記憶部30と、この入力情報記憶部30に接続された予測モデル構築部10と、この予測モデル構築部10に接続されたリスクスコア計算部20と、このリスクスコア計算部20へ接続された進行中プロジェクト情報記憶部40とを有する。影響因子候補情報記憶部60は第3の記憶手段に相当し、影響因子情報決定部70は第4の処理に相当する。

10

## 【 0 1 5 7 】

リスク評価装置3は、第1の実施の形態に係るリスク評価装置1と基本的に同様の構成要素を有しており、影響因子候補情報記憶部60と影響因子情報決定部70とをさらに有している点が異なっている。

## 【 0 1 5 8 】

第3の実施の形態に係るリスク評価装置3の構成要素のうち、第1の実施の形態に係るリスク評価装置1の構成要素と同一のものは、同一の参照符号を付してそれら構成要素の詳細な説明は省略する。

20

## 【 0 1 5 9 】

## [ 3 . 1 . 1 . 影響因子候補情報記憶部 ]

影響因子候補情報記憶部60は、影響因子候補情報1200を記憶する機能を有する。影響因子候補情報1200は、影響因子及びその区分情報の候補の集合である。リスク評価装置3のオペレーター、ユーザ、管理者などが影響因子及びその区分情報の候補を複数入力した情報である影響因子候補情報1200に基づいて、影響因子候補情報記憶部60は、各影響因子と区分情報について期待区分情報リフト値を算出し、この期待区分情報リフト値に基づいて影響因子及びその区分情報の候補を比較し、影響因子情報400に含める影響因子及びその区分情報を決定する。なお、タイミング情報は、影響因子及びその区分情報の決定では必要としないが、予め影響因子候補情報1200に含められていてもよい。

30

## 【 0 1 6 0 】

## [ 3 . 1 . 2 . 影響因子情報決定部 ]

影響因子情報決定部70は、過去プロジェクト情報の符号化を行い、期待区分情報リフト値を計算し、期待区分情報リフト値に基づいて影響因子情報400に含める影響因子及びその区分情報を選択し、選択した影響因子及びその区分情報からなる影響因子情報400を生成し、入力情報記憶部30に記憶させる機能を有する。

## 【 0 1 6 1 】

40

## [ 3 . 1 . 2 . 1 . 過去プロジェクト情報の符号化 ]

影響因子情報決定部70は、過去プロジェクト情報200の符号化を行い、符号化過去プロジェクト情報を生成する。過去プロジェクト情報200を符号化する方法は、第1の実施の形態と同様である。ただし、符号化で使用する区分情報は、影響因子候補情報1200のものを用いる。

## 【 0 1 6 2 】

## [ 3 . 1 . 2 . 2 . 期待区分情報リフト値の算出 ]

影響因子情報決定部70は、影響因子候補情報1200に含まれる影響因子情報の候補（影響因子  $X_i$ 、区分情報  $k = \{0, 1, \dots, n\}$ ）のそれぞれに対して、下記式6で定義される期待区分情報リフト値  $EL_{X_i}$  を計算する。

50

【 0 1 6 3 】

【 数 1 3 】

$$EL\_Xi = \frac{1}{N} \sum_k n_k |L_{Xi,k}| \dots \text{(式6)}$$

【 0 1 6 4 】

但し式6において

【 0 1 6 5 】

【 数 1 4 】

10

$$N = \sum_k n_k \quad n_k : \text{区分情報 } k \text{ の数}$$

【 0 1 6 6 】

【 数 1 5 】

$$|L_{Xi,k}| = \begin{cases} L_{Xi,k} & (L_{Xi,k} \geq 1) \\ \frac{1}{L_{Xi,k}} & (L_{Xi,k} < 1) \end{cases}$$

20

但し、 $L_{Xi,k}$  : 影響因子  $X_i$ 、区分値  $k$  の区分情報リフト値

【 0 1 6 7 】

影響因子情報決定部 70 は、算出した期待区分情報リフト値  $EL\_Xi$  を参照して、影響因子候補情報 1200 に含まれる影響因子の内、どの影響因子を影響因子情報 400 に含めるかを決定し、決定した影響因子を含む影響因子情報 400 を生成し、これを入力情報記憶部 30 に記憶させる。

【 0 1 6 8 】

なお、複数の影響因子情報の候補がある場合には、期待区分情報リフト値  $EL\_Xi$  が高いものを選ぶ方が、プロジェクト成功 / 失敗の予測精度が高まる可能性が大きいと考えられるが、必ずしも期待区分情報リフト値  $EL\_Xi$  が高いものから順に選ぶ方式に本実施の形態は限定されるものではない。

30

【 0 1 6 9 】

期待区分情報リフト値の算出の具体例を示す。同一の影響因子  $X_i$  について異なる区分情報を有する、2つの影響因子の候補について期待区分情報リフト値  $EL\_Xi$  を算出し、比較する。

【 0 1 7 0 】

(1) 第1の候補

第1の候補は、影響因子  $X_1$ 、区分情報 {0:[MIN, 2.5), 1:[2.5, MAX], NA:欠損値} という内容を有する。

40

【 0 1 7 1 】

この第1の候補の期待区分情報リフト値  $EL\_Xi$  を算出する。但し、過去プロジェクト情報 200 は、図2に示すデータを使用したものとする。

【 0 1 7 2 】

影響因子情報決定部 70 は、過去プロジェクト情報 200 及び第1の候補の区分情報から符号化過去プロジェクト情報及び区分情報リフト値を生成し、この符号化過去プロジェクト情報及び区分情報リフト値から以下の値を得る。

【 0 1 7 3 】

【数 1 6】

$$n_0 = 5, \quad n_1 = 4, \quad N = \sum_{k=0,1} n_k = 9$$

【0 1 7 4】

【数 1 7】

$$L_{X_1,0} = 0.7333 \rightarrow |L_{X_1,0}| = \frac{1}{0.7333} = 1.3637$$

10

【0 1 7 5】

【数 1 8】

$$L_{X_1,1} = 1.32 \rightarrow |L_{X_1,1}| = 1.32$$

【0 1 7 6】

影響因子情報決定部 7 0 は、これらの値を上記式 6 にあてはめ、第 1 の候補の期待区分情報リフト値  $EL_{X_i}$  を以下のように得る。

【0 1 7 7】

20

【数 1 9】

$$EL_{X_1} = \frac{5 \times 1.3637 + 4 \times 1.32}{9} = 1.3443$$

【0 1 7 8】

(2) 第 2 の候補

第 2 の候補は、影響因子  $X_1$ 、区分情報 {0:[MIN, 2.0), 1:[2.0, MAX], NA:欠損値} のときという内容を有する。第 1 の候補とは同一の影響因子であるが区分情報の内容は異なっている。

30

【0 1 7 9】

この第 2 の候補の期待区分情報リフト値  $EL_{X_i}$  を算出する。但し、第 1 の候補の場合と同様に過去プロジェクト情報 2 0 0 は、図 2 に示すデータを使用したものとする。

【0 1 8 0】

影響因子情報決定部 7 0 は、過去プロジェクト情報 2 0 0 及び第 2 の候補の区分情報から符号化過去プロジェクト情報及び区分情報リフト値を生成し、この符号化過去プロジェクト情報及び区分情報リフト値から以下の値を得る。

【0 1 8 1】

【数 2 0】

40

$$n_0 = 3, \quad n_1 = 6, \quad N = \sum_{k=0,1} n_k = 9$$

【0 1 8 2】

【数 2 1】

$$L_{X_1,0} = 0.55 \rightarrow |L_{X_1,0}| = \frac{1}{0.55} = 1.8182$$

【0 1 8 3】

50

【数 2 2】

$$L_{X_1,1} = 1.2571 \rightarrow |L_{X_1,1}| = 1.2571$$

【0 1 8 4】

影響因子情報決定部 7 0 は、これらの値を上記式 6 にあてはめ、第 2 の候補の期待区分情報リフト値  $EL_{Xi}$  を以下のように得る。

【0 1 8 5】

【数 2 3】

$$EL_{X_1} = \frac{3 \times 1.8182 + 6 \times 1.2571}{9} = 1.4441$$

10

【0 1 8 6】

上記第 1 及び第 2 の影響因子情報の候補の期待区分情報リフト値  $EL_{Xi}$  を比較すると、第 2 の候補の期待区分情報リフト値  $EL_{Xi}$  (=1.3443) よりも、第 2 の候補及び期待区分情報リフト値  $EL_{Xi}$  (=1.4441) の方が若干高い。よって、影響因子情報 4 0 0 の決定において、影響因子情報決定部 7 0 は、第 2 の候補を採用し、影響因子情報 4 0 0 に第 2 の候補のデータを含める。これによりプロジェクト成功 / 失敗の予測精度が高まる可能性がある。

20

【0 1 8 7】

[ 3 . 2 . 第 3 の実施の形態の動作例 ]

図 1 6 に示したリスク評価装置 3 の動作例を説明する。図 1 7 に第 3 の実施の形態に係るリスク評価装置 3 の主たる動作の例を示したフローチャートを示す。

【0 1 8 8】

まずリスク評価装置 3、より詳しくは影響因子情報決定部 7 0 は、過去プロジェクト情報 2 0 0、プロジェクト失敗定義情報 3 0 0、影響因子候補情報 1 2 0 0 に基づいて符号化過去プロジェクト情報を生成する (S210)。

【0 1 8 9】

次にリスク評価装置 3、より詳しくは影響因子情報決定部 7 0 は、ステップ S210 で生成した符号化過去プロジェクト情報から、影響因子候補情報 1 2 0 0 に含まれる影響因子の候補のそれぞれについて期待区分情報リフト値  $EL_{Xi}$  を算出する (S220)。

30

【0 1 9 0】

次にリスク評価装置 3、より詳しくは影響因子情報決定部 7 0 は、期待区分情報リフト値  $EL_{Xi}$  に基づいて影響因子情報 4 0 0 に含める影響因子及びその区分情報を決定し、決定した影響因子及びその区分情報を含む新たな影響因子情報 4 0 0 を生成し、これを入力情報記憶部 3 0 に記憶させる (S230)。

【0 1 9 1】

次にリスク評価装置 3 はリスクスコア算出処理を実行する (S240)。リスクスコア算出処理は、第 1 の実施の形態におけるリスクスコア算出処理 (図 1 2、ステップ S10 ~ S40 参照) と同一の処理であるので、処理内容の詳述は省略する。なお、リスクスコア算出処理 (S240) で使用される影響因子情報 4 0 0 は、ステップ S230 で生成、記憶された影響因子情報 4 0 0 である。

40

以上で第 3 の実施の形態に係るリスク評価装置の動作の説明を終了する。

【0 1 9 2】

[ 3 . 3 . 第 3 の実施の形態の利点 ]

第 1 の実施の形態の予測モデル構築において、ユーザやオペレータ等によって行われる影響因子の選定や区分情報の設定などの作業に最も手間がかかり、その良し悪しが最終的なプロジェクト失敗の予測精度にも大きく影響する。

【0 1 9 3】

50

第3の実施の形態では、期待区分情報リフト値を評価関数として、最適な影響因子の選定、及び、区分情報の設定を行うことが可能となることにより、予測モデル構築作業の手間を大幅に低減すると共に、プロジェクト失敗の予測精度向上も期待できる。

【0194】

[4. 第4の実施の形態]

第1、第2、第3の実施の形態を組みあわせてもリスク評価装置は成立する。第1、第2、第3の実施の形態を組みあわせたリスク評価装置4の構成例を示す機能ブロック図を図18に掲げる。なお、第1、第2、第3の実施の形態と同一の構成要素については同一の参照符号を付す。

【0195】

[5. まとめ、その他]

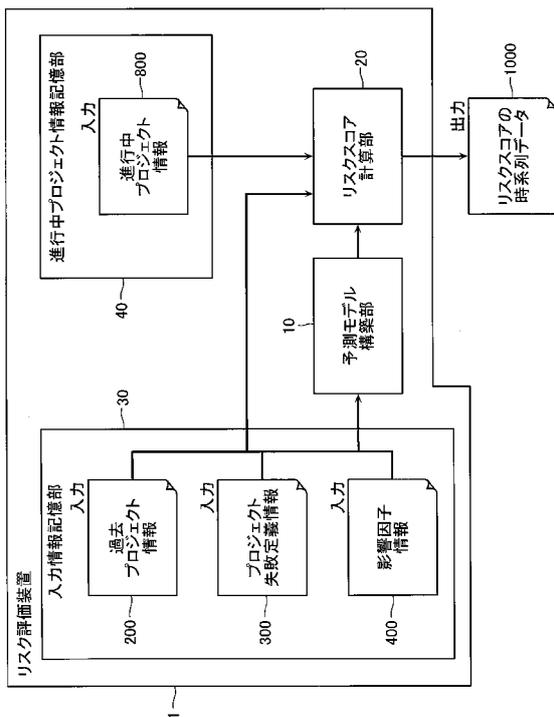
以上、本発明の実施の形態を説明したが、本発明はこれらに限定されるものではなく、発明の趣旨を逸脱しない範囲内において、種々の変更、追加、組み合わせ等が可能である。

【符号の説明】

【0196】

1、2、3、4・・・リスク評価装置； 10・・・予測モデル構築部； 20・・・リスクスコア計算部； 30・・・入力情報記憶部； 40・・・進行中プロジェクト情報記憶部； 50・・・プロジェクト失敗確率計算部； 60・・・影響因子候補情報記憶部； 70・・・影響因子情報決定部

【図1】



【図2】

	211	212	213	214	215	216	
	プロジェクト名	仕様レビュー指摘密度	設計レビュー指摘密度	設計レビュー効率	単体テストバグ密度	結合テストバグ密度	出荷後バグ密度
201	PJ1	0.5	1.4	2.1	欠損値	1.1	1.2
201	PJ2	5.7	0.1	3.6	2.9	2.5	2.1
201	PJ3	欠損値	1.3	0.8	1.4	1.7	1.3
201	PJ4	2.1	0.7	3.3	1.8	0.3	0.9
201	PJ5	3.6	4.5	1.2	欠損値	3.2	4.8
201	PJ6	1.2	0.5	3.1	7.5	欠損値	3.2
201	PJ7	2.6	1.1	11.3	5.2	2.1	10.5
201	PJ8	4.2	0.9	0.1	0.9	0.8	1.5
201	PJ9	1.9	1.3	欠損値	12.5	2.3	4.8
201	PJ10	0.8	0.2	2.1	2.6	2.8	4.1

【図3】

	411	412	413
	影響因子	区分情報	タイミング情報
401	仕様レビュー指摘密度	[MIN, 2.5], [2.5, MAX], 欠損値	T = 1 (要求仕様化工程)
401	設計レビュー指摘密度	[MIN, 1.2], [1.2, MAX], 欠損値	T = 2 (設計工程)
401	設計レビュー効率	[MIN, 3.5], [3.5, MAX], 欠損値	T = 2 (設計工程)
401	単体テストバグ密度	[MIN, 5.5], [5.5, MAX], 欠損値	T = 3 (単体テスト工程)
401	結合テストバグ密度	[MIN, 2.2], [2.2, MAX], 欠損値	T = 4 (結合テスト工程)

【図4】

	511	512	513	514	515	516
	X1	X2a	X2b	X3	X4	Y(結果)
501	PJ1	0	1	0	NA	0
501	PJ2	1	0	1	0	1
501	PJ3	NA	1	0	0	0
501	PJ4	0	0	1	0	0
501	PJ5	1	1	0	NA	1
501	PJ6	0	0	0	1	NA
501	PJ7	1	0	1	0	0
501	PJ8	1	0	0	0	0
501	PJ9	0	1	NA	1	1
501	PJ10	0	0	0	0	1

【図5】

	Y = 0 (成功)	Y = 1 (失敗)
$X_i = k$	c1	c3
$X_i \neq k$ かつ $X_i \neq NA$	c2	c4

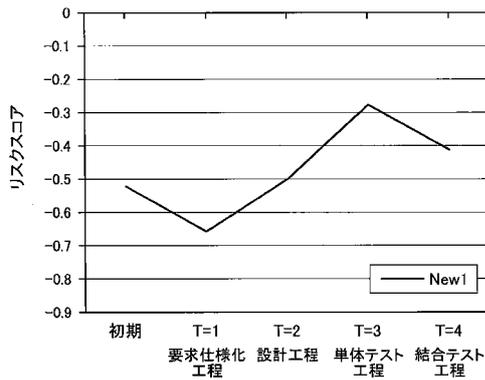
【図6】

	Y = 0 (成功)	Y = 1 (失敗)
$X_1 = 0$	4	1
$X_1 \neq 0$ かつ $X_1 \neq NA$	2	2

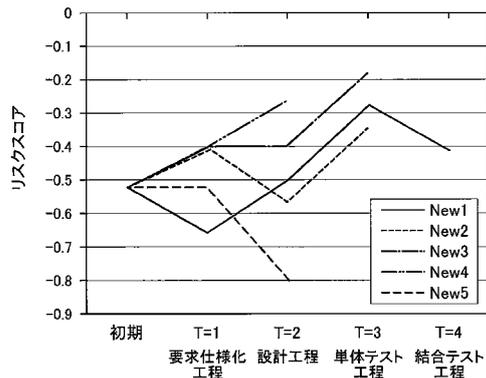
【図7】

	Y = 0 (成功)	Y = 1 (失敗)
$X_1 = 1$	2	2
$X_1 \neq 1$ かつ $X_1 \neq NA$	4	1

【図10】



【図11】



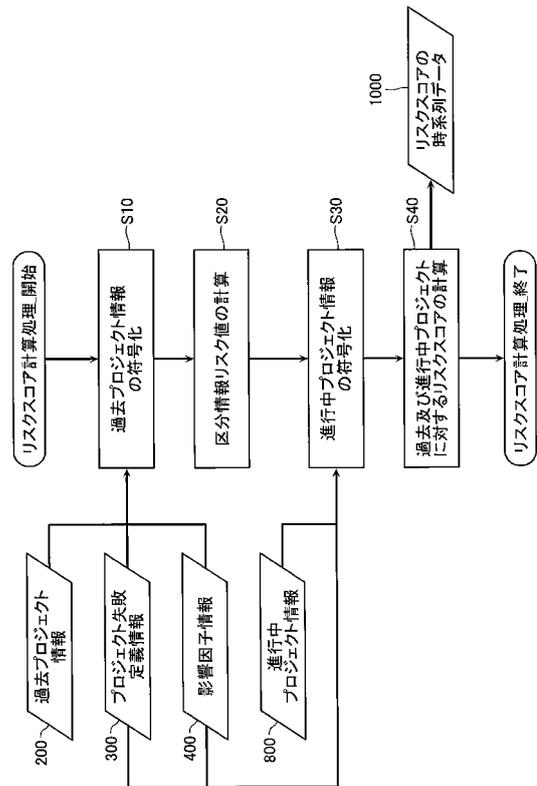
【図8】

	811	812	813	814	815	816
	プロジェクト名	仕様レビュー指摘密度	設計レビュー指摘密度	設計レビュー効率	単体テストバグ密度	結合テストバグ密度
801	New1	1.2	1.5	欠損値	7.9	1.2
801	New2	2.7	0.8	欠損値	6.2	未実施
801	New3	3.2	欠損値	欠損値	11.8	未実施
801	New4	4.5	欠損値	4.5	未実施	未実施
801	New5	欠損値	0.3	2.1	未実施	未実施

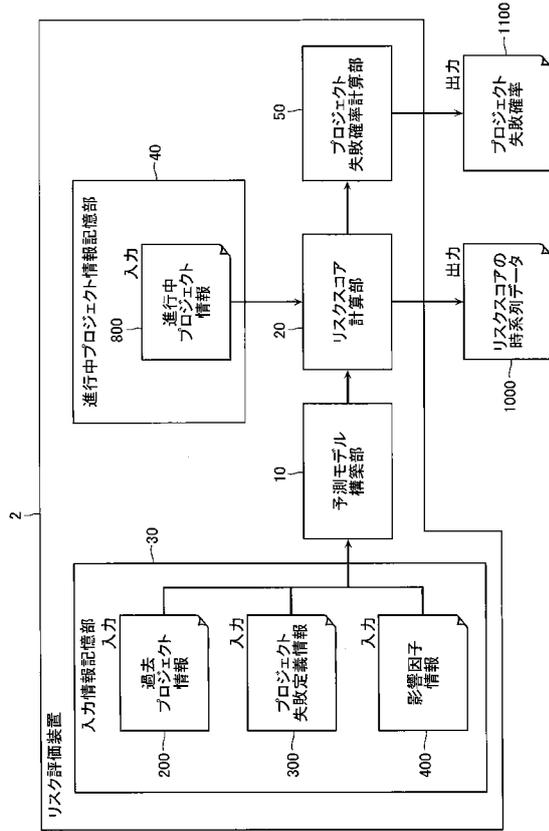
【図9】

	911	912	913	914	915	916
	X1	X2a	X2b	X3	X4	Y(結果)
901	New1	0	1	NA	1	0
901	New2	1	0	NA	1	未実施
901	New3	1	NA	NA	1	未実施
901	New4	1	NA	1	未実施	未実施
901	New5	NA	0	0	未実施	未実施

【図12】



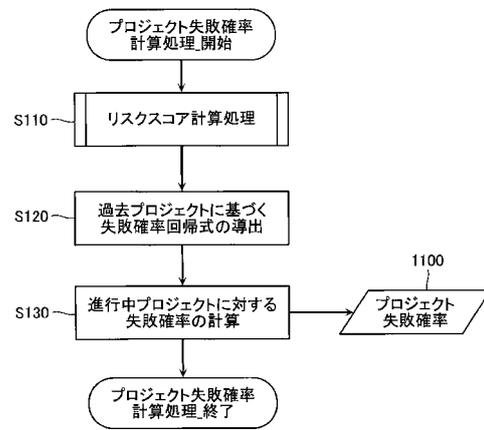
【図 1 3】



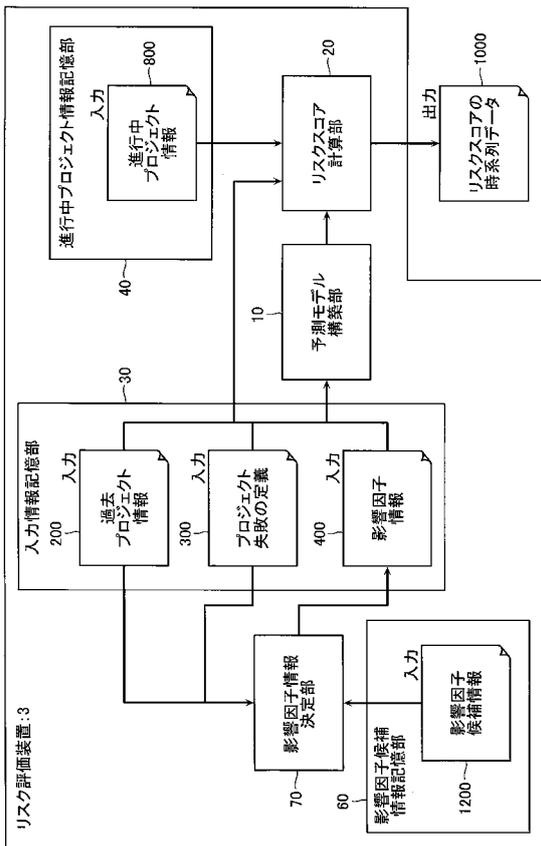
【図 1 4】

	PJ1	PJ2	PJ3	PJ4	PJ5	PJ6	PJ7	PJ8	PJ9	PJ10
Y	0	0	0	0	1	0	1	0	1	0
X	-0.7387	-0.4534	-0.7501	-0.9640	-0.2281	-0.7043	-0.7087	-0.9517	-0.1568	-0.9517

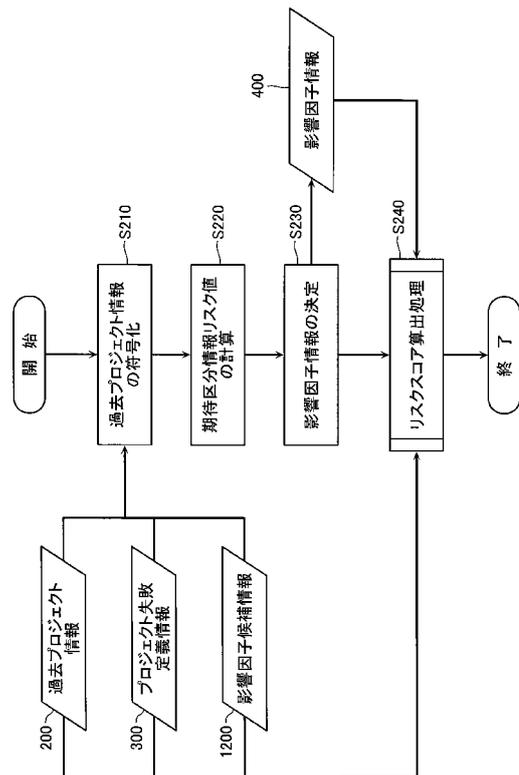
【図 1 5】



【図 1 6】



【図 1 7】



【図 18】

