

サブテーマ名：共同研究全般にわたるデータベース構築・管理 小テーマ名：サブテーマ4の3 データマイニングツールの開発
サブテーマリーダー（所属、役職、氏名）千葉県産業振興センター（研究員）甲賀弘、村上 雅利 小テーマリーダー研究従事者（所属、役職、氏名）（株）数理システム 水田 千益 研究従事者（所属、役職、氏名）（株）数理システム（共同研究員）徐 良為
<p>研究の概要、新規性及び目標</p> <p>①研究の概要</p> <p>かずさDNA研究所においては、その保持するマウス長鎖cDNAライブラリーをもとに、マイクロアレイを利用して、遺伝子発現情報を収集することになっている。本収集データは、構築予定のマイクロアレイ・データベースシステムに格納される。そのデータは膨大なものになり、その解析には、規格化、フィルタリング、クラスタリングなどのデータマイニングの諸手法が必要となる。また、マイクロアレイ解析に特有な手法も必要である。本テーマの研究では、数理システム開発の汎用データマイニングツールを用い、マイクロアレイデータの解析に必要な手法の確立、およびそれをもとにしたマイクロアレイ専用データマイニングツールの開発を目指して来た。</p> <p>②研究の独自性・新規性</p> <p>統計及びデータマイニング手法を用いて、マイクロアレイ解析を行う単発的な先行事例があるが、データマイニングと統計のあらゆる過程も統一的に対応可能な分析環境を提供する例が少なく、本ツールの含まれる独自開発、分析精度の向上及びデータ可視化に大きく有効と言われる機能として、ネットワーククラスタリング、サポートベクトルマシン、One-Class SVM、クラスアソシエーション分析などが含まれている。</p> <p>③研究の目標（フェーズ毎に数値目標等をあげ、具体的に）</p> <p>フェーズⅠ（H14-H15）：</p> <p>データマイニングの環境開発、マイニング前処理機能を設計及び実装を行う。具体的には、</p> <p>分析環境： データマイニングの過程を処理フローで表し、データの入出力、プロジェクト保存、インポート・エクスポートを一元管理するシステムを構築する。</p> <p>前処理： マイクロアレイを用いて、cDNAライブラリーに保持されるマウス長鎖cDNAライブラリーから得られる発現情報をテーブル形式にまとめ、不要な情報の削除（変数選択）、欠損補填、規格化、基本情報集計などを行うに必要な機能を実装する。</p> <p>フェーズⅡ（H16-H18）：</p> <p>データマイニング分析ツールの開発、主に：</p> <ul style="list-style-type: none"> ・ アソシエーション分析、多階層アソシエーションルール抽出機能、時系列付きアソシエーション分析、ユーザ対話型の結論からルールの抽出機能及びWebダイアグラム機能の設計、実装、テストを行いました。 ・ 新分類ツール Support Vector Machine の設計と実装 ・ 新分類ツール Naïve Bays の設計と実装 ・ 新クラスタリングツール One Class SVM の設計と実装 ・ クラスアソシエーション分析機能の設計と実装 <p>フェーズⅢ（H18. 11.25事業完了日以降）：</p> <p>主にデータマイニングの新機能開発及び、バイオ全般への適用性検証を行った。具体的には、</p> <ul style="list-style-type: none"> ・ ルールベース予測ツールの開発 ・ ネットワークグラフ、クラスタリング機能の実装 ・ データ可視化ツール：データビューア、データハンドリングの実装 ・ Visual Mining Studio を用いて、KDD-2001のバイオデータ分析を行い、報告書にまとめた。

研究の進め方及び進捗状況（目標と対比して）
当初の目標値にほぼ達成した。

主な成果
具体的な成果内容：
汎用データマイニングツール Visual Mining Studio がリリースしたことによって、広範囲のユーザに好評を頂き、特にバイオ業界にも注目を集めつつあります。
特許件数：0 論文数：0 口頭発表件数：2

研究成果に関する評価
1 国内外における水準との対比
日本発の汎用データマイニングツールであり、他社製品と比較しては、機能が豊富で、使いやすい、実装速度速いとの評価を頂いた。

2 実用化に向けた波及効果
バイオ業界だけでなく、様々の分野への適用に使われている。

残された課題と対応方針について
今後、よりバイオに特化したツールを作成したいである。

	J S T負担分（千円）							地域負担分（千円）							合 計
	H 13	H 14	H 15	H 16	H 17	H 18	小計	H 13	H 14	H 15	H 16	H 17	H 18	小計	
人件費									260	1,980	1,930	2,175		6,345	6,345
設備費									308						308
その他研究費 (消耗品費、 材料費等)															
旅費															
その他															
小 計									568	1,980	1,930	2,175		6,653	6,653

代表的な設備名と仕様 [既存（事業開始前）の設備含む]
J S T負担による設備：なし
地域負担による設備：コンピューターシステムなど

※複数の研究課題に共通した経費については按分する。