

平成26年度北海道大学情報基盤センター共同研究成果報告書

1. 研究領域番号 A3

2. 研究課題名 発見的特徴把握のための情報縮約・変数選択・クラスタリングの研究

3. 研究期間 平成 26 年 4 月 1 日 ～ 平成 27 年 3 月 31 日

4. 研究代表者

氏 名	所属機関・部局名	職 名	備 考
黒田 正博	岡山理科大学・総合情報学部	教授	

5. 研究分担者

氏 名	所属機関・部局名	職 名	備 考
森 裕一	岡山理科大学・総合情報学部	教授	
榊原 道夫	岡山理科大学・総合情報学部	教授	
飯塚 誠也	岡山大学・アドミッションセンター	教授	
久保田 貴文	多摩大学・経営情報学部	准教授	
南 弘征	北海道大学・情報基盤センター	准教授	
桜井 裕仁	大学入試センター 研究開発部・試験基盤設計研究部門	准教授	
片山 浩子	岡山理科大学大学院数理・環境システム専攻(博士課程[後期])	大学院生	

6. 共同研究の成果

下欄には、当該研究期間内に実施した共同研究の成果について、その具体的内容、意義、重要性等を、共同研究申請書に記載した「研究目的」と「研究計画・方法」に照らし、800字～1,000字で、できるだけ分かりやすく記載願います。文章の他に、研究成果を端的に表す図表を貼り付けても構いません。なお、研究成果の論文・学会発表等を行った実績（発表等の予定を含む。）があれば、あわせて記載して下さい。

本研究は、マーケティング、環境、Web、ゲノム等の分野で得られるビッグデータに対する効率的なデータの特徴把握に関するものであり、

- 1) 先行研究等の情報収集と分析・整理
- 2) 次元縮約手法（変数選択手法を含む）の検討
- 3) あらゆる尺度を考慮した情報縮約とクラスタリングの研究
- 4) 新しい融合手法の開発
- 5) 計算環境の開発および効率化の検討

をテーマに掲げ、データの複雑さを考慮したあらゆる尺度のデータを処理できる情報縮約・クラスタリング融合手法を整備・提案するとともに、発見的情報表現が可能なインタフェースと高速な計算環境を提供することを目的とした。

個別の研究打ち合わせの後、同じA3に属する共同研究課題「ビッグデータ解析のためのシンボリックデータ解析法の研究」のメンバーと合同で次の研究集会を行った。

■ テーマ（研究課題名）：

「発見的特徴把握のための情報縮約・変数選択・クラスタリングの研究」
「ビッグデータ解析のためのシンボリックデータ解析法の研究」

■ 日 時：2014 年 11 月 28 日（金） 10:00～15:00

■ 場 所：北海道大学情報基盤センター 北館 4 階会議室

■ プログラム

10:00～10:10 水田正弘（北海道大学・情報基盤センター）
開会の挨拶

10:10～10:40 清水信夫，中野純司（統計数理研究所），山本由和（徳島文理大学）
カテゴリー変数を含む集約的シンボリックデータのクラスタリング

10:40～11:10 *片山琴絵，山口類，井元清哉，宮野悟（東京大学・医科学研究所）
伝統的医学データの統計解析について

11:10～11:30 総合討論

11:00～13:00 昼食

13:00～13:30 桜井 裕仁（大学入試センター・研究開発部）
大学入学志願者の基礎的学力測定のための「数理分析力」試験の開発と検討

13:30～14:00 黒田 正博（岡山理科大学・総合情報学部）
混合モデルにおける EM アルゴリズムの初期値選択について

14:00～14:30 五十嵐 千人，松井 佑介，南 弘征，水田正弘（北海道大学）
相関を考慮した分布値データの解析法について

14:30～15:00 総合討論

本研究会では，データが有する特徴を把握するための解析手法の提案および統計数値計算法の改良，実データへの適用例などの研究成果が示され，さらに参加者全員で数理分析力の能力を評価する方法に関する議論もおこなった．上記の研究発表のうち，桜井による発表が目標 1)と 4)，黒田の発表が目標 3)から 5)，そして五十嵐らによる発表が目標 2)と 3)の研究テーマに関係する成果となる．