

プログラム名：脳情報の可視化と制御による活力溢れる生活の実現

PM名：山川 義徳

プロジェクト名：脳ビッグデータ

委 託 研 究 開 発

実 施 状 況 報 告 書 (成 果)

平 成 2 7 年 度

研究開発課題名：

脳アンチエイジング

研究開発機関名：

株式会社アラヤブレインイメージング

研究開発責任者

金井 良太

I 当該年度における計画と成果

1. 当該年度の担当研究開発課題の目標と計画

平成27年度における研究目標としては、次のような数値目標をもって開始した。「健常者の心理的特徴を $R^2 > 0.5$ で予測できる技術の実現を目標とする。これを一般向けの個人特性予測システムとして公開を実現し、少なくとも5つ以上の心理指標についての予測モデルを提供する。またユーザからのフィードバックを元に、インターフェイスの改善を行う。また、本研究開発の成果を海外での学会で発表を行う。(年次計画書より引用)」。すなわち、主に2種類の開発に分類可能で、ひとつは、解析ツールの普及にむけた脳情報解析の自動化・簡易化にむけた統合的プラットフォームの実装である。もうひとつは、T1強調画像のような脳の構造画像を元に機械学習を用いて、個人の心理特性の予測モデル構築の構築である。具体的な成果目標として、1. 「健常者の心理的特徴を $R^2 > 0.5$ で予測できる技術の実現」、2. 「個人特性予測システムの構築」、3. 「5つ以上の心理指標についての予測モデルの構築」を目標として研究開発に取り組んだ。これらの予測を実現する脳データとしてはT1強調画像をもとにVoxel-Based Morphometry (VBM)の前処理を行った3次元データを入力画像として、心理特性が出力されるシステムの構築を目指した。

2. 当該年度の担当研究開発課題の進捗状況と成果

2-1 進捗状況

脳MRI構造画像情報は3次元のボクセルデータであり、1mm解像度の場合、数10万の数値データからなる。個人データの数に比べ、ボクセルデータの数が過剰であるため、心理指標について一般化線形モデルによる予測モデルを構築するには、効率的にその自由度を削減し、汎化したモデルを構築することが必要となる。このような問題に対する手段として、elastic net と呼ばれる正則化の手法を用いた(e.g. Friedman et al, 2007, 2010)。特に今回はglmnet と呼ばれるパッケージを用いてMATLAB上で開発を行った(Qian et al., 2013)。Elastic net では、一般化線形モデルにL1 ノルムおよびL2 ノルムによる拘束項を付加して、さらに各展開係数のうち影響の小さいものを0に落とすことにより、逐次的に自由度削減を行った。

2-2 成果

VBMの前処理により、1.5mm刻み、8mm半値幅でスムージングされた灰白質画像(c1)を用いて個人データの推定を行った。灰白質そのものの大きさに個人差や性別さがあるため、ボクセルデータを、灰白質と白質の体積の合計で規格化した。きわめて小さい値によりモデルが不安定になるのを防ぐために、ボクセルデータをThreshold(=0.1)により小さな値を除外した。前処理において8mmでスムージングをかけていることから、xyz各空間座標に対して5グリッドずつデータを単純間引きしても独立性は失われないと考え、これにより3000程度まで自由度を削減した。この間引いた3000自由度の脳画像ボクセルデータをもとに、個人データを予測するモデルをelastic net で構築したが、この状態でもなお自由度は過剰な状態であり、elastic net の自由度削減の性質は有効に機能すると考えられる。過学習がアルゴリズム上の問題でないことを確認するために、手動でnolds=10のnested cross validationを行った。すなわち、10グループのうちの9組を教師データとしてモデルを構築し、残りの1組をテストデータとして予測を行うというプロセスを計10回繰り返し、実測値と予測値をプロットした(図1)。

以上の手順に基づいて、HCP511人分のデータについて、性別、BIG5についての予測モデル構築を試みた。用いたパラメーターは $\alpha=1$ とし、 λ は上記手順により項目ごとに独立に決定した。図1ではBig

Five の Extraversion という外向性の心理特性の解析結果を示したが、他の項目もほぼ同様の精度での予測が実現できた。

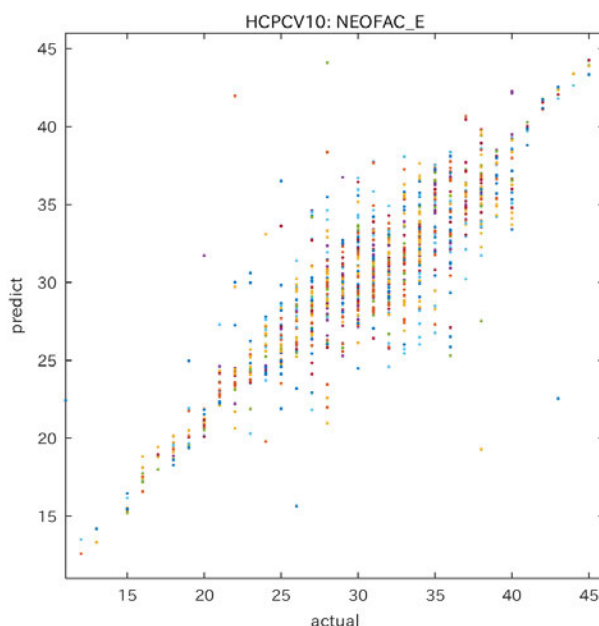


図 1 Big5 の Extraversion について Nested CV を行った結果。横軸が実測、縦軸が予測値。10 個のグループについて 10 回の予測モデル構築が行われた結果を重ねてプロットした。色の違いはグループの違いを示す。

2-3 新たな課題など

本研究開発により実現した、Elastic net による自由度削減を用いた予測モデル構築は良好に機能しており、予測値と実測値の関係を確認した。しかし、これらのモデルが本当に過学習に落ちていないかの評価、他のモデルに比した優劣の評価をどのように行うかについての課題は残る。すなわち、汎化性能の評価が課題である。汎化性能の評価のために、手持ちのデータセットを分割して交差検証を行うというアプローチをとることができるが、脳のボクセルの次元(数 10 万)に比べて、データ数が 500 人分程度と非常に少ないため、汎化の評価を行うためにデータを半分に分けようとする、そもそもそのデータの分け方に引きずられ、本来のデータの能力を引き出すことができなくなる。elastic net は本質的に縮小推定の枠組みに属するものであり、予測されるテストデータが、訓練データと同じモデル分布に従っているかどうかにより、汎化の評価が引きずられるという課題が残る。データ数を増やす仕組みを作ることが、今後のモデルの精度向上および予測モデルの汎化性能評価のために欠かせないと考えられる。

3. アウトリーチ活動報告

本研究活動の一環として、VBM での手法を他の MRI 画像モダリティである拡散 MRI 画像への拡張を試み、計算に時間のかかる NODDI などの新しい解析手法の高速化と自動化についての準備を行ってきた。その過程で実現した NODDI 導入に必要な前処理のノウハウと高速化の可能性についてのチュートリアルを実施し、また資料の公開を行った。