戦略的創造研究推進事業 CREST
研究領域「イノベーション創発に資する
人工知能基盤技術の創出と統合化」
研究課題「自然言語処理による心の病の理解:
未病で精神疾患を防ぐ」

研究終了報告書

研究期間 2016年 12月～2019年 3月

研究代表者:岸本 泰士郎
(慶應義塾大学医学部
精神・神経科学教室 専任講師)

§１ 研究実施の概要

（1）実施概要

精神疾患は罹患率が高く罹病期間も長いため、人々の生活の質（QOL)を低下せしめるものとして重要な疾患群である。世界保健機関（WHO）などが行う世界疾病負担（Global burden of disease）調査によると、うつ病、不安症、統合失調症、躁うつ病、薬物依存などを含む精神疾患は、障害生存年数（years lived with disability, YLD）において他の医学領域を押さえ第一位、22.9%を占める。ところが、血液所見や画像所見などの客観的評価が利用可能な他の身体疾患と異なり、精神疾患は生物学的な指標に乏しく、診断の不一致、治療開始基準の不明確さなど、様々な問題につながっている。

そこで、本プロジェクトでは精神科医が診断を行う際に最も重要視する「患者の言葉」に表れる疾患の特徴量の抽出や定量を行う、新しい精神科研究領域の創出に取り組んだ。具体的には精神医学の専門的知見と自然言語処理を融合させることで、代表的な精神疾患の言語的な特徴を同定・定量することを試みた。本プロジェクトで開発される技術によって、疾患の前兆や早期発見、あるいは鑑別が困難な症状を分離できるようになる。さらに疾患の前駆期や早期での症状の発現のタイミングを捉えることで、効果的な予防・治療技術の開発が可能になる。精神科医療現場での使用のみならず、個人ユーザによる自動診断機、企業でのメンタルヘルス対策、認知症予防などの社会サービスに用いることができ、社会的価値は計り知れない。

本プロジェクトの実施に際しては、正確な診断によって紐づけされた患者の言語データを慶應義塾大学の研究代表者（岸本泰士郎）のグループ（岸本 G）が収集し、分担者（狩野芳伸）のグループの静岡大学（狩野 G）が自然言語処理を用いて最適な特徴量抽出の設計と診断結果の学習を行うという研究体制を築いた。スモールフェーズでは上述の目的のために以下の 2 つの Study を実施した。

Study 1 においては、国際的な診断基準を満たすうつ病・双極性障害、統合失調症、不安症、認知症患者、あるいは健常者を対象として、心理士が対話形式のインタビューを行い、音声データを収集した。各被検者は 1 回 60 分程度の自然会話を心理士と行い、用意された種々のテーマに沿った会話を行った。また、インタビューの度に、国際的に用いられている評価尺度を用いて、重症度評価を行った。収集したデータは、音響学的な解析、人手を用いた文字起こしを行ってテキストデータに変換、さらには自動音声認識を用いたデータに変換し、自然言語処理を用いて機械学習による分類を試みた。

Study 2 においては、同様に国際的な診断基準を満たすうつ病・双極性障害、統合失調症、不安症、認知症患者を対象に、ツイッター等の SNS データを研究グループに共有していただき、同様の解析を行った。

2019年 3 月末の時点で、Study 1 において 289 時間程度の音声データを、また、Study 2においては512 万字程度のテキストデータの収集を行うことができた。11 月までに得られたデータを用いて、自然言語処理および機械学習を用いた疾患の同定精度の検証を行ったところ、高確率での罹患有無の推定に成功した。特に、うつ病、認知症、不安症においては、個人ユーザによる自己診断ツールなどの実用化や、医療機器開発を想定したコア技術としても実現性の高い結果を得ることができており、本技術のさらなる社会的応用も期待できる結果と考えている（例：会社全体のムードを検出・改善する技術、認知機能の低下した高齢者に対する社会サービス等）。

（2）顕著な成果

＜優れた基礎研究としての成果＞
1.
概要：
　うつ病・双極性障害、統合失調症、不安症、認知症患者の会話コーパスを作成した。本コーパスは国際的な診断基準を満たす精緻な診断や疾患の重症度と紐づけられている。3 月末で登録数 156 名、レーティングセット 293 セットとなり、およそ 293 時間程度分のデータを収集しており、研究を継続することで 400 時間を超える音声コーパスになることが明らかである。疾患情報に紐づけられた日本語の会話コーパスは他に類を見ず、国際的にもここまで精緻な診断・重症度情報が付与されたコーパスはなく非常に貴重なデータとなる。

2.
概要：
　上記会話コーパス等を用いて、精神疾患の有無やその重症度を推定する自然言語処理を用いた解析アルゴリズムを構築した。既に高精度の結果が得られており、このような方法での識別が可能であることを示した。このことは、精神医学の臨床診断の在り方や研究の方向性にも影響を与える結果である。国際的に類似の研究成果も一部報告されているが、それを上回る結果が得られている。

＜科学技術イノベーションに大きく寄与する成果＞
1.
概要：
　うつ病の診断やその重症度の推定が可能な音響学的特徴や自然言語処理を用いた解析モデルの構築を行った。一定の精度が得られており、本技術の応用から、うつ病の早期サインや再発の予兆を捉える自己診断サービス、治療への反応性を定量する医療機器への応用のみならず、会社全体のムードを定量しうつ病等による休職を防ぐサービス、社会全体の気分を定量し政策に生かすサービスなど、多くの事業に展開である。

2.
概要：
　認知症の推定が可能な音響学的特徴や自然言語処理を用いた解析モデルの構築を行った。本技術の応用ら、認知症の早期サインを捉える自己診断サービス、重症度を定量する医療機器への応用のみならず、認知機能が低下される場合に財産を保護するような金融サービス等、多くの事業に展開可能である。

3.
概要：
　不安症（パニック、社交不安障害、強迫性障害など）の推定が可能な音響学的特徴や自然言語処理を用いた解析モデルの構築を行った。不安症（それぞれ人口の 1-2%の有病率）を抱える患者の治療効果判定ツールや、認知行動療法に用いるフィードバックツールとしての事業展開が可能である。

＜代表的な論文＞
M. Sakishita, T. Kishimoto, A. Takinami, Y. Eguchi, Y. Kano. Large-scale Dialog Corpus Towards Automatic Mental Disease Diagnosis, International Workshop on Health Intelligence (W3PHAI 2019), AAAI-19 (2019) 2019/1/27

M. Sakishita, C. Ogawa, K. J. Tsuchiya, T. Iwabuchi, T. Kishimoto, Y. Kano. Autism Spectrum

Disorder's Severity Prediction Model Using Utterance Features for Automatic Diagnosis Support, International Workshop on Health Intelligence (W3PHAI 2019), AAAI-19 (2019) 2019/1/27

Matsumoto T, Sutoh C, Asano K, Seki Y, Urao Y, Yokoo M, Takanashi R, Yoshida T, Tanaka M, Noguchi R, Nagata S, Ohshiro K, Numata N, Hirose M, Yoshimura K, Nagai K, Nakagawa A, Sato Y, Shimizu E, Kishimoto T. Internet-based cognitive behavioral therapy with real-time therapist support via videoconference for patients with obsessive-compulsive disorder, panic disorder, and social anxietry disorder: A pilot single-arm trial. Journal of Medical Internet Research. In press. 2018

§２ 研究実施体制

（1）研究チームの体制について

① 慶應義塾大学グループ
研究代表者：岸本 泰士郎（慶應義塾大学医学部精神・神経科学教室 専任講師）
・データ収集：プロトコール（臨床研究実施計画書）策定、倫理指針・個人情報保護の観点からの検討、倫理申請、被験者リクルート・データ収集
・プラットフォーム構築：静岡大学グループと共同した自動データ収集システムの実装
・データ解析：精神科専門家の見地から解析を支援

② 静岡大学グループ
研究分担者：狩野 芳伸（静岡大学情報学部自然言語処理 准教授）
・データ収集：必要データ属性・フォーマットの策定
・プラットフォーム構築：自動データ収集システムの実装
・データ解析：特徴量アノテーション、特徴量抽出ツール、および自動診断システム開発を目指した試作段階の解析アルゴリズムの構築

（2）国内外の研究者や産業界等との連携によるネットワーク形成の状況について

理化学研究所・革新知能統合研究センター（AIP）との技術交流（人的な技術交流も含む）、浜松医科大学、志學館大学と連携、ならびに企業との今後のデータ収集や実装場所の提供などに関して、継続的な議論の場を持つことにより強固なネットワーク形成をした。