

2022 年度研究助成（2 年助成） 研究実績報告書

代表研究者	井口 元三
研究テーマ	機械学習および深層学習を活用したメンタルヘルス関連労働災害の予測モデル開発

I. 研究の背景と概要

【背景】厚生労働省の「患者調査」によれば、職場におけるメンタルヘルス不調者数が増加しており、過労死、うつ病、自殺へと繋がる事例も増えています。この問題に対処するため、「ストレスチェック」制度が導入され、全労働者に義務付けられました。現在、この制度は約 83%の実施率を誇り、最も広く行われているメンタルヘルス調査制度です。しかし、総合点でストレス反応を評価するこの制度では、重要な項目が埋もれてしまう可能性が指摘され、新たな予測モデルの必要性が高まっています。

【目的】本研究の目的は、これまでに報告してきた結果をもとに、機械学習および深層学習（ディープラーニング）を活用して、ストレスチェックに基づくメンタルヘルス不調者を見逃さないための新しい予測モデルを構築することです。

【方法】2019～2020 年度に、本学的全職員を対象にストレスチェックを実施しました。そのデータを元に、機械学習や深層学習（ディープラーニング）を用いて予測モデルを試みました。ストレスチェックの各項目を前処理し、メンタル相談の有無をターゲットにして、メンタルヘルス不調者の予測モデルを作成しました。

II. 研究の成果

【結果】基礎統計データとしてのストレスチェック対象者 5,315 人の中で、回答者は 3,352 人（回答率 63.1%）でした。その中でメンタル相談受診者は 116 人（3.5%）でした。得られたデータセットを教師データ：予測データの 8：2 に分割し、教師データを用いて各種機械学習アルゴリズム（ランダムフォレスト、ロジスティック回帰、サポートベクターマシン（SVM）等）により複数の予測モデルを作成しました。ハイパーパラメータチューニングおよびデータセットを 10 分割して交差検証を行い、正解率（Accuracy）が 95%を超える複数の予測モデルが作成されました。

機械学習におけるベストモデルはこれまでの報告と同様に Light Gradient Boosting Machine (Light GBM) であり、Accuracy: 0.9654, AUC: 0.7395 でした。さらに、詳細に解析したところ Top5 のモデルは、Light GBM、Random Forest Classifier (0.9648, 0.7344)、Quadratic Discriminant Analysis (0.9648, 0.5000)、Extra Trees Classifier (0.9648, 0.7712)、CatBoost Classifier (0.9638, 0.7708) でした。

また、得られた各予測モデル（弱学習器）を集約したアンサンブル学習を行うことで、高精度な予測モデルを構築しました。Top5 の予測モデルを用いたスタッキングでは (0.9503, 0.6399) とベストモデルの Light GBM に及びませんでしたが、Light GBM を用いた Bagging (0.9648, 0.7539)、Boosting (0.9643, 0.7741) では良好な予測モデルが構築できました。

さらに、今回の研究では機械学習だけでなく深層学習（ディープラーニング）ソフト Neural Network Console を用いて、より効率の良い予測モデルを求めて構造探索を行い、その結果、予測精度の高いニューラルネットワークによる予測モデル（Accuracy: 0.9632～0.9876）を複数構築することに成功しました。これは、これまでの報告よりも精度が高くなっていました。

今後、より大規模なストレスチェックデータを用いてより精度の高い予測モデルを構築することを目指します。

【結語】多重的に関わる職業性ストレスにおけるメンタルヘルス不調者予測モデルの構築に成功しました。将来的には、疾病を予防する質の高い保健サービスの創出が期待されます。

この成果は、効果的なストレス管理やメンタルヘルスの向上に向けた重要な一歩であり、社会的な健康増進に貢献するものと期待されます。