

(続紙 1)

京都大学	博士 (経済学)	氏名	毛 柏林
論文題目	Application of Machine Learning in Stock Prediction, Portfolio Optimization and Experimental Investigation of People's Behavior towards AI Stock Prediction (株式予測とポートフォリオ最適化のための機械学習応用および人工知能の株式予測に対する人間行動の実験研究)		
(論文内容の要旨)			
<p>従来の金融論では数理モデルを導出するために市場と投資者に理想的な仮説を設けながら理論を発展させてきた。他方、過去数十年にわたって仮説と現実の差異により起きた大規模なファンド破綻と金融危機が問題視されている。その反省に立ち、これからの金融論は社会の発展に伴い、理論での市場仮説をビッグデータ時代の社会情報と関連する機械学習技術を利用して検証することと、その結果を踏まえてより現実と相応しい理論方法を提案することが必要である。一方、現実に見られる投資者の合理・非合理的な意思決定に対して、行動ファイナンスのアプローチにより証拠に基づく実験研究を行うことも不可欠である。本博士論文では、機械学習の手法を利用した株式価格の予測とポートフォリオの最適化を行う実証研究を行うとともに、投資者のロボットアドバイザーに対する依存度を調べる実験研究を行ったものである。第一章は全体をまとめた要旨である。</p> <p>第二章は金融理論の基盤となる「効率的市場仮説」(セミストロング型)に対する検証の一助となる研究を行った。この研究では、フォロワーが多いツイッターアカウントから得た株式に関する膨大なツイート文章と深層学習モデルを用いて、株価日足終値の動き(上昇と下落)を予測している。この研究に利用する深層学習モデルとしては、一般的な自然言語処理において優れた性能を有する複数のモデルを使い、ツイート情報の抽出を行っているが、その結果効率的市場仮説と整合的であることが示されている。</p> <p>第三章では株式収益率の平均・分散をコントロールしながらポートフォリオを決定する現代ポートフォリオ理論の代替手法について考察されている。現代ポートフォリオ理論では分散をリスクの指標とするのが標準的であるが、投資リスクは収益率の分散自体ではなく、投資者が損する確率であることが常に指摘されている一方、多くの実証研究で「ハイリスク・ハイリターン」に反する事例も世界各地の金融市場にて観察されているため、この研究は平均・分散トレードオフの分析手法に代わり、「密度」をコントロールしながらポートフォリオを決定する代替手法を考える。具体的には、オリジナルポートフォリオからクラスタリング法の一つであるDBSCANを利用する提案手法の「密度最適化」処理で、「最適な密度」を持つ部分集合を抽出する。パフォーマンス指標としてシャープ・レシオとジェンセンのアルファを用いて、この部分集合がオリジナルポートフォリオのパフォーマンスを上回ることを示し、新規に提案する「密度」に基づくコントロール手法は平均・分散トレードオフの分析手</p>			

法の有望な代替手法であることを示している。

第四章では、投資家がロボットアドバイザーの投資助言を合理的に参考とすることができるか否かを経済実験により考察している。この実験の設定は、参加者が過去一年での株価日足終値のプロットを見て、ロボットアドバイザーの予測についての助言を参考にしながら、1ヶ月先の株価日足終値を予測し、最終予測精度によって金銭的報酬を得るものである。更に、パフォーマンス水準の異なる二つのロボットアドバイザーを用意する。性能の良いロボット（人の平均パフォーマンスを上回る）と悪いロボット（人の平均パフォーマンスを下回る）に被験者を半分ずつ割り当て、一方のグループには前者のロボット、他方には後者が与えられる。更に、各グループ内で練習ステージがない組、練習ステージでの自身パフォーマンスが見える組、練習ステージでの自身パフォーマンスとロボットパフォーマンスが見える組という三つの組に分けてトータルで6つの群を作って上述の実験を行っている。その結果、練習ステージがない組の実験参加者は、ロボットへの依存度がロボットのパフォーマンス水準により影響されないことが示された。他にも、練習ステージがある組の実験参加者のうち、自身のパフォーマンスを一度見た参加者のロボット依存度は、練習ステージのない参加者のロボット依存度を顕著に下回ることなど、いくつかの興味深い結果を得ている。

(続紙 2)

(論文審査の結果の要旨)

第二章ではツイッターの文章が株価の上下動に対して説明力があるかどうかを調べたものである。「効率的市場仮説」(セミストロング型)は、すべての公開している情報が既にトレーダーに利用され、トレーダーの売買行為を通じて金融資産の価額に反応されているため、ツイート文章により株価の動きを予測することができないことを示唆する。この研究の結果によると、深層学習モデルで抽出されたツイート情報を利用する株価予測の精度は、当て推量(Random Guess)の予測精度と大して差がなく、どちらも50%ほどである。従って、深層学習モデルによって引き出されたツイート情報は日次の頻度のデータについては「効率的市場仮説」と整合的であることが明らかになった。独自の大規模ツイッターデータと最新のテキストデータに関する機械学習の手法を用いた分析であり、十分に興味深い結果を得ている。

第三章では、DBSCANを利用する「密度最適化」処理によるポートフォリオの部分集合を取り出して新たなポートフォリオを作ったときのパフォーマンスを、過去のアメリカの株価の実データを用いて調べている。パフォーマンス測度をシャープ・レシオに設定したとき、特徴が異なる各オリジナルポートフォリオに対して、「密度最適化」処理によるポートフォリオはほぼ一律にパフォーマンスを向上させるという結果が得られ、それ以外の標準的な三つのベンチマーク手法よりもパフォーマンスを上回った。他方、パフォーマンス測度をジェンセンのアルファとしたときは、「密度最適化」処理は必ずしもパフォーマンス向上に繋がらなかった。そこで、アルファを指標とする際に対応するために「部分的密度最適化」処理を提案し、これを用いるとオリジナルポートフォリオのパフォーマンスを向上させる結果が得られることを示した。以上の結果から、「密度」に基づく分析手法は平均・分散の分析手法の有望な代替手法であることが明らかになった。これにより、最もよく用いられる平均・分散のトレードオフに依拠したポートフォリオ構築以外の新しい有望な手法が提供され、クラスタリング技術とポートフォリオ選択という現実課題に対する実装にも有益な研究となっている。ただし、今後の課題としては他のデータへ拡張した際の外的妥当性をどのように確認、確保するかを考える必要があり、これは更なる研究がまたれるところである。

第四章では、パフォーマンスの異なる株価予測ロボットへの依存度を調べる経済実験を行っている。高性能と低性能のロボットを与えられた時に、いくつかの設定でロボットへの依存度がどのように変わるかをRCTによって調べており、行動ファイナンスのテーマとしても極めて独創的な研究であると高く評価できる。今回の被験者は大学の学生に偏りがあるため、第三章の結果と同様に外的妥当性についての検証が必要であろう。しかし、この研究も株価予測ロボット活用の実装に向けての示唆に富む内容であり、学術研究としても実務的にも重要な貢献であることは明らかである。

よって、本論文は博士(経済学)の学位論文として価値あるものと認める。また、令和5年1月18日、論文内容とそれに関連した事項について試問を行った結果、合格と認めた。

要旨公表可能日： 年 月 日以降