

## 【学位論文審査の要旨】

### 1. 研究の背景と研究目的

サイコロのように理想的に反復できる事象や、堅牢な物理法則に従って何度でも繰り返せる物理実験などと違い、我々が観測できる金融現象はひとつの歴史に沿った事象だけである。実証実験でバックテストをする場合でもその束縛から逃れることはできず、その意味で我々は「歴史」というバイアスのかかったデータのみを使って実験していることになる。近年金融でも広く使われるようになってきた機械学習技術においても、その訓練データが実証データである限り、学習結果は歴史のバイアスを受け、その分、将来の事象に対する堅牢性を損なうことになる。

この問題を解決するために、観測した（経験した）歴史以外の時系列を人工的に発生させデータセットを拡大する、「データ拡張」と呼ばれる手法が近年注目を浴びてきている。拡張データを用いて訓練をすれば、歴史のバイアスの軛（くびき）を逃れ、より堅牢で汎化性のある学習結果を得ることが期待できる。ただここで問題になるのは、拡張データの「リアリティ」である。人工的に生成されたデータは何らかの意味でこのリアリティが担保されなくてはならない。その解答のひとつとして注目されている技術が敵対生成ネットワーク（GAN）である。これは **generator** と **discriminator** という2つのニューラルネットワークを切磋琢磨させて最終的にリアリティのあるデータを生成させる仕組みである。

本論文は、GAN を核とした機械学習技術を使ってリアリティのある株価のヒストリカル・データを人工的に生成し、そうして得られた拡張データによって訓練することでより堅牢な学習効果を得ようという試みである。

### 2. 本論文の構成

本論文は英語で記述され全 5 章で構成されている（本文 76 ページ）。各章の見出しは以下の通り。

#### Chapter 1. Global Introduction

##### 1.1 Introduction

##### 1.2 Research motivation

##### 1.3 Purpose of research

##### 1.4 Our contribution

##### 1.5 Structure of thesis

#### Chapter 2. Data Augmentation of High Frequency Financial Data Using Generative Adversarial Network

##### 2.1 Introduction

##### 2.2 Prerequisite

2.3 Method

2.4 Data

2.5 Result

2.6 Conclusion

## Chapter 3. Discrete Signature and its Application to Finance

3.1 Introduction

3.2 Discrete Signature

3.3 An implementation of discrete signature

3.4 An application of discrete signature to finance

3.5 Concluding remarks

## Chapter 4. Stock Market Simulation by Micro-Macro GAN

4.1 Introduction

4.2 Related work

4.3 Prerequisite

4.4 Method

4.5 Data

4.6 Result

4.7 Conclusion

## Chapter 5. Global conclusion and discussion

Bibliography

Index

なお、本論文の第2章は査読付き学術誌“IEEE Xplore”第24巻（2021年6月号）に掲載された成富氏が第1著者の論文“Data Augmentation of High Frequency Financial Data Using Generative Adversarial Network”（共著者：足立高德）を、第3章は2021年12月に arXiv.org にポストした成富氏が第2著者の論文 "Discrete Signature and its Application to Finance"（arXiv:2112.09342 [q-fin.MF]）（共著者：足立高德）を、また第4章は、2023年6月に熊本市で開催される第37回人工知能学会全国大会に投稿中の成富氏が第1著者の論文“Stock Market Simulation by Micro-Macro GAN”（共著者：足立高德）を基礎に、加筆修正したものである。

### 3. 本論文の要旨

第1章では、全体的な解説と歴史の中での本研究の位置づけを行っている。

第2章では、最初に注文データのマイクロなダイナミクスに着目し、Wasserstein GAN を用いた注文データ生成とデータ拡張の有効性について考察している。具体的には、WGAN を使って生成した人工板から価格時系列データを作りデータ拡張を行った後、個別銘柄の価格予測精度が高まったことを確認した。

第3章では、マクロなダイナミクスである価格変動に着目し、有効な特徴量を得るための手法として高次元曲線の特徴量を抽出する際に近年よく使われるようになった signature の利用方法を検討した。特に、金融時系列データに適用するために離散的なデータから直接 signature を計算する手法である離散 signature を提案し、金融時系列データへの応用について議論した。

第4章では、受注データのマイクロダイナミクスに対する WGAN と価格のマクロダイナミクスに対する signature を用いた損失関数を提案し、マイクロとマクロのダイナミクスを結合したデータ生成手法として Micro-Macro GAN を提案し、その有効性を議論した。具体的には、第2章で開発したデータ生成手法に加えて、マクロな価格データに第3章で開発した離散 signature を用いて生成した人工データを加え、新たに「Micro-Macro GAN」と名付けた手法を開発した。この新手法で生成した注文データは現実のデータと似た分布を持ち、その価格ダイナミクスも現実のデータで観測される長期周期のあるダイナミクスを持つことが確認できた。

第5章では、全体の議論と結論について述べている。

#### 【論文審査結果の要旨】

##### 1. 審査結果

21世紀の現在は深層学習に代表される機械学習の技術が広い分野に浸透してきている。多くの事例では大量のセンサーによって得られたビッグデータとその解析技術という形で革新が進んでいる。金融現象への機械学習の応用でも例外ではなく、計算機シミュレーションを使って頑強な予測を行うには大量なデータがかかせない。しかし物理現象と違い基本的にひとつの歴史しか観測できない金融現象では十分な量のデータを用意することが難しいことが多々ある。この問題を解決するため、近年では現実的な人工データを自動生成し（「データ拡張」）、こうして得られた拡張データセットによって機械学習を行う手法が注目を集めている。本論文では株式市場における高頻度取引に着目し深層学習モデルを用いてトレーダーが発注する注文の人工的な時系列の生成を敵対生成ネットワーク(GAN)を用い

て試みている。

本論文では、最初に、Wasserstein GAN を使って生成した人工板から価格時系列データを作りデータ拡張を行った後、個別銘柄の価格予測精度が高まったことを確認した。次に、高次元曲線の特徴量を抽出する際に近年よく使われるようになった **signature** を離散データから直接作る技術を開発し、それを用いてマクロな価格データに人工データを加え、新たに **Micro-Macro GAN** と名付けた手法を開発した。この新手法で生成した注文データは現実のデータと同程度の分布を持ち、その価格ダイナミクスも現実のデータと同程度の長期周期のあるダイナミクスを持つことを確認し、人工市場シミュレーションを実現するために必要なリアリティのあるデータを得られたとした。このようにして構築した人工市場は、様々な時間地平でのデータ拡張やバックテスト、シナリオテスト、自動売買アルゴリズムの改善などへの利用が期待される。

一方、本論文では、計算スピードの高速化、より統計的精密さを持った結果の検討、**signature** のパラメタ選択の検討など、いくつかの改善余地が考えられるが、まったく新しいデータ拡張の手法を提案し実装してみせたことは、金融リスク管理を含む重要な応用への道を開いたという意味で、本プログラムの博士論文の基準を満たしていると考えられる。

## 2. 合否判定

本審査委員会は、学位申請者である成富佑輔氏に対して、令和4年12月22日に本論文について公開審査を実施した。その結果、申請者が博士学位を取得するにふさわしい学識を有していることが確認できた。よって、本審査委員会は申請者成富佑輔氏に対して、東京都立大学・博士(経営学)の学位を授与することが適当であると判定する。