

LSTMを用いたトレーディングカードゲーム中古市場における 時系列分析

Time Series Analysis of Used Trading Card Game Market Using Long Short-Term Memory

宍戸 建元[†] 新美 礼彦^{††}

[†] 公立はこだて未来大学大学院 システム情報科学研究科 〒041-8655 北海道函館市亀田中野町 116 番地 2

^{††} 公立はこだて未来大学 システム情報科学部 〒041-8655 北海道函館市亀田中野町 116 番地 2

E-mail: [†]{g2121025,niimi}@fun.ac.jp

あらまし 近年、中古販売市場は急速に発展している。その中でも、特にトレーディングカードゲーム (以下、TCG) 市場の成長は著しい。その TCG 市場において、ゲーム内での性能によって価格が 10 倍以上に急高騰する例が存在する。また価格も日ごとに更新され、株価のような時系列な変化が確認されている。しかし、その事象について詳細に研究された例はわずかであり、時系列的なデータとして TCG の中古価格を扱った研究は少ない。そこで本研究では、TCG の 1 つである「Magic: The Gathering」における中古販売市場を対象とし、時系列解析を適用することで価格の高騰予測を目指す。特に、LSTM(Long Short-Term Memory) を利用した価格予測を行うことで、時系列的な変化に妥当性があるか検証する。また、実際の価格データを利用した学習を行い、その予測結果から高騰の規則性や季節性などを考察する。

キーワード データマイニング, 機械学習, 価格予測, LSTM, 時系列分析

1 はじめに

1.1 背景

近年、中古販売市場は急速に発展している。特に、EC サイトやフリマアプリを利用した中古販売が盛んになり、情報技術を活用した中古商品の売買も増加している。『令和 2 年度 産業経済研究委託事業 (電子商取引に関する市場調査) 報告書』[1] では、令和 2 年度の CtoC-EC 市場規模が 1 兆 9,586 億円、前年度比 12.5 % 増と拡大傾向にあることを論じている。また、拡大の背景に感染症による外出自粛の呼びかけや EC の利用が推奨されたことがあると論じた。さらにライフスタイルの変化により在宅時間が増え、カテゴリ構成がエンタメ・ホビー用品等インドア向けにシフトしたことも論じている。

その中で、特にトレーディングカードゲーム (以下、TCG) の中古販売取引が大きく成長しつつある。イーベイ・ジャパン社のレポート [2] では、越境 EC の 7~9 月期における TCG のカテゴリ成長率が前年比の 237 % と増加し、4 期連続 3 桁成長と市場規模が現在も拡大が続いていると論じている。また、現在では一部カードの希少価値から美術品としての需要も高まっており、鑑定が行われた希少なカードが 1,000 万円以上の価格で取引される例も報告されている。

1.2 TCG 市場における急高騰について

この TCG 市場において、既存中古製品の需要が突然高まり、

価格が急騰する現象が存在する。図 1 は、実際の価格高騰の例である。この例について、2019/04/01 時点までは商品 1 枚がトリム平均価格 500 円程度で取引されている。しかし、この商品と相性の良い新製品の情報が公開されたことで価格が高騰し、最終的に 05/16 時点で 9,456 円と、18 倍程度まで上昇している。

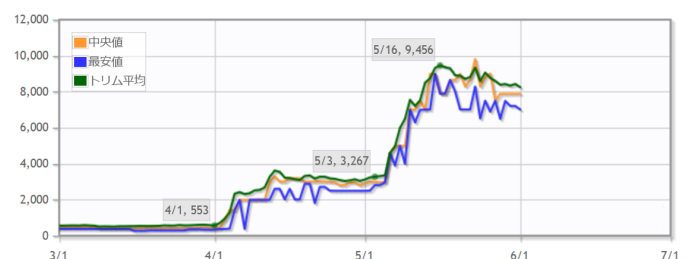


図 1 実際の商品 [3] の価格高騰の例

このように、通常は安価な価格で取引されていた商品が、様々な要因により需要が高まり、最終的に大きく高騰する例が存在する。また近年では、メーカーから商品情報が事前に公開されることや、EC サイトの普及により中古売買市場でも素早い価格更新が可能になったことから、日ごとにその価格が変化し、急騰の速度も上昇している。

しかし TCG 市場の高騰現象について、機械的に長期間の予測を行った例は少ない。その理由として TCG 商品の性能は文字情報を多く含んでいるが、形式的な記述でないものが多く、

機械的に相性を推測し記述することが困難な点があげられる。また、相性のいい新商品の発売だけでなく、競技大会の結果、インフルエンサーの発言など、高騰の理由が多岐にわたる点も挙げられる。そのため、実店舗では勘と経験によって商品の値付けをしており、現在でもその高騰を予測することは難しい。

1.3 目的

前述の背景を踏まえ本研究では、TCG のタイトルである「Magic:The Gathering」[4] における中古販売市場を対象とし、商品の高騰現象について実際の販売価格履歴データを用いた機械学習による予測を目指す。

本研究では特に、その価格が日ごとに変化していく様子から価格の時系列的な変化に着目し、LSTM ネットワークを用いた時系列分析による予測を目指す。

またその学習についても複数のパラメータを変更した学習を行い、予測精度の最適化と TCG の価格予測として LSTM を用いることが可能か議論していく。

2 関連研究

2.1 穴戸らの研究

TCG 市場に関する近年の関連研究として、穴戸ら[5]の研究が挙げられる。この研究では新商品発売前後の高騰に着目し、高騰有無の2クラス分類問題として決定木学習を行うことで、高騰要因の簡易な分析を行った。その結果定性的な価格の場合と同様に、高騰要因として希少度を示す属性が有意であることを示した。また、新商品の発売時期ごとの高騰要因の違いとして、ゲーム中の性能を示す属性があることを論じている。

穴戸らの研究は、対象が TCG 市場の中古価格に対する時系列分析である点は本研究と一致している。そのため、今回の実験に利用する価格データセットの作成は、穴戸らの研究を参考に行った。また対象とする価格帯も、穴戸らと同じように低価格帯や高価格帯による影響を省くため、本研究においても範囲を限定して分析する。

しかし、穴戸らは決定木学習を用いたクラス分類による分析であり、本研究で用いる LSTM による実価格予測とは異なる。また、利用する価格データの対象期間も異なっており、穴戸らは新商品発売前後2週間の範囲を取り扱っているが、本研究は約10ヶ月分のデータに対し、期間中全日の価格から学習する。

2.2 Matthew らの研究

TCG 市場の予測に関する関連研究として、Matthew ら[6]の研究があげられる。この研究では、TCG 市場の価格変化を4種類に分類し、クラス分類問題として中古販売価格の要因分析と予測を行った。またこの研究の特色として、過去の価格を学習の属性に利用している点がある。現在から6日前以内までの価格、価格差、使用率、販売量の差、をそれぞれ求め、そこにマナコスト、特定のルール上で使えなくなるまでの日数、を加えた属性を学習に用いた。その結果、価格変化の要因として現在から3日以内の価格・使用率・販売量が最も有効な属性であると論じている。

本研究との違いとして、Matthew らは学習手法が回帰分析と SVM であり、本研究の目指す LSTM による分析とは異なる。また高騰に対しクラス分類問題として適用している点も、実価格の予測を目指す本研究とは異なっている。

しかし、数日前の価格を利用して予測する点は本研究の目指す分析手法と一致しているため、この研究の手法を参考にし、過去6日分の価格を利用し予測を行った。

2.3 LSTM に対する関連研究

LSTM に関連する研究として、小松ら[7]の研究や白澤ら[8]の研究があげられる。小松らの研究では電力使用量の予測に LSTM を用い、気温、降水量、天気等を組み合わせた複数入力からの学習を行った。また白澤らの研究では、クラスタ分析によって学習データの選定を行い、その結果から LSTM を用いた鹿児島市の1時間後の降水量の予測を行った。これらの研究は本研究と取り扱う分野は別であるが、LSTM による予測を行う点は本研究と一致しているため参考にした。

さらに、LSTM を用いた市場の分析として、宮崎ら[9]の研究や、仲岡ら[10]の研究が挙げられる。宮崎らの研究では、株価予測に対する Convolutional Neural Network(CNN) の応用として、LSTM を用いた手法に一定の効果があることを示した。また仲岡らの研究では、日経平均株価の翌日の終値を予測する問題として LSTM を利用し、移動平均も含めた価格の予測を行った。本研究はこれらの研究と違い株価市場を取り扱うものではないが、時系列的に価格が変化する市場に機械学習を用いた先行例であり、本研究の類似例であるため、これらの文献を参考に今回の学習手法として LSTM を用いることを決定した。

3 提案する分析手法

本研究で提案する分析手法として、価格が日ごとに更新され時系列的な変化をしていることから、時系列分析を用いた実価格予測を提案する。

特に、価格への短期的な影響だけでなく長期的な影響も考慮するため、今回は LSTM を用いた実価格の予測を提案する。これは前述の参考研究において発売時期による変化が確認された点、属性として6日前の価格の履歴が利用されている点、実際の高騰要因が外部要因なども含め多岐に渡ることから、分析の手法として妥当だと判断し、LSTM による分析を行った。

学習に用いるデータについても、商品の性能に関する情報は用いず、株価予測のように価格履歴のみを利用した予測を目指す。また、TCG 中古市場の価格範囲が数十円から数百万円と価格の幅が広いと、超低価格帯や高価格帯の影響を受けやすい。そのため、価格履歴データの範囲についても、穴戸らの研究を参考に1USD~150USD までの価格帯に対する予測を提案する。

予測の方針については、Matthew らの研究や株価に関する先行研究を参考に、過去6日分の価格から次の日の価格を予測する問題として学習を適用する。ただし、今回は既存商品の価

格予測を扱う問題として学習を行うものとする。これは、期間中に発売した商品について、発売までの価格の取り扱いを含めて複雑になると考え、今回の実験では考慮しないものとする。

4 実験

本章では、提案した分析手法を用いて LSTM による中古販売価格の予測を試みる。また、学習に用いるパラメータを一部変化させ予測精度の最適化を行い、その結果から価格予測に用いる手法として妥当か否かを確認する。

4.1 中古販売価格データセットの取得

実験に使用した中古販売価格データセットについて、「Magic: The Gathering」の海外 EC サイトにおける販売価格情報が集約された “MTG JSON” [11] というサイトから中古販売価格データを取得し作成を行った。このサイトでは EC サイト 3 店舗分の販売価格データセットが存在する。今回はそのうち最も欠損日の少なかった EC サイト 1 店舗 (“Card Kingdom” [12]) における中古販売価格データを取得し利用した。また、データセットの欠損した日の価格については補完を行わず、そのまま次の日のデータとして分析に用いた。

実験で使用した商品の中古販売価格データセットは、2021/01/04～同年 12/10 時点までの、価格データを取得することが可能であった 296 日分のデータを対象とした。また対象とする商品は、ゲーム上の公式ルールで使用不可能なカードや foil 等の特殊加工のカードを除き、2021/01/04 以前までに販売開始された商品のみを扱うものとした。

実際に取得したデータのうち、一部商品は価格を取得した際の通貨単位が異なっており、USD と EUD の 2 種類が存在した。そのため今回は、1USD=1.2EUD として価格の変換を行い通貨単位を統一することとした。その後、超低価格帯や高価格帯の影響を考慮しデータ範囲を絞るため、期間内の価格の最小値が 1USD 未満、最大値が 150USD より大きい商品を削除した。

最終的に今回取得した中古販売価格データセットは、商品データ 7,338 件 × 296 日分の価格履歴データ、総計 2,172,048 件を利用した。

4.2 学習用データセットの作成

先述のように取得した中古販売価格データセットを加工し、LSTM に用いるための学習用データセットを作成した。学習には、6 日分の過去データから次の日の価格予測を行う問題のため、予測したい次の日の価格 p_t を目的変数、6 日前までの価格履歴 $\{p_{t-6}, p_{t-5}, p_{t-4}, p_{t-3}, p_{t-2}, p_{t-1}\}$ を従属変数 (属性) とするデータセットを作成した。

最終的に 296 日分のデータから、1 商品につき 6 日前分までの価格履歴 290 組、これを対象期間までに発売された商品 7,338 件に対して行い、2,135,358 件の学習用データセットを作成し実験を行った。

4.3 LSTM による学習について

本研究で利用した LSTM として、TensorFlow の Keras [13] ライブラリを使用し学習を行った。パラメータの設定について、今回は隠れ層の数を 500 個とし、バッチサイズを 32、勾配降下法の最適化アルゴリズムに Adam を用い、学習率を 0.001、損失関数には平均二乗誤差 (mean_squared_error) を利用した。さらに、活性化関数として線形関数 (linear) を利用し、交差検証用に 1 割のデータを利用し学習を行った。また、epoch 数は今回 3,10 の 2 種類の場合を別々に学習し、それぞれ epoch 数の指定による違いを確認した。

実際に学習に使用したデータは、学習用データセットのうちの 8 割である 1,708,287 件を利用し、残りの 2 割である 427,071 件を評価用のテストデータとして分割し、実験を行った。

実験後に、学習中の loss (損失)、val_loss (交差検証時の損失) の推移を確認し、LSTM による学習が問題なく行われたか、また最適な epoch 数に関する考察を行った。

最終的に構築された学習器の精度を確かめるため、テストデータに対して予測を行い、平均二乗誤差 (MSE)、二乗平均平方根誤差 (RMSE)、決定係数 (R^2) による評価を行った。その結果から予測精度を判断し、TCG の価格予測として LSTM を用いることが可能か考察した。

さらに、実際の商品に対しても価格の予測ができていないか確認するため、テストデータのうち期間中の変化量 (=最大値 ÷ 最小値) が最も大きい商品上位 3 件に対して、価格予測を個別に行った。その予測結果と実際の価格履歴とのあてはまり具合から、予測精度についての考察した。

5 結果

LSTM による学習の結果について、各ステップにおける loss と val_loss の推移を、図 2、図 3、図 4、図 5 に示す。

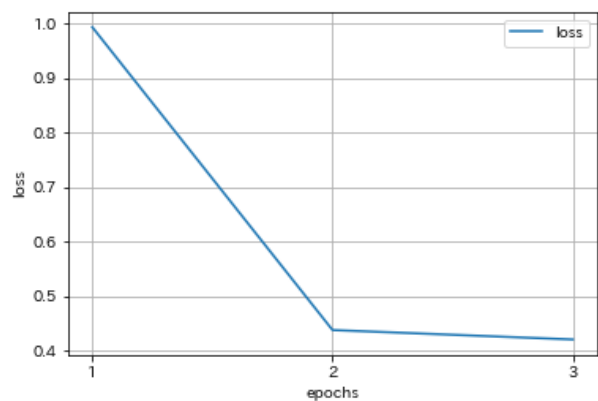


図 2 各ステップごとの loss (epoch 数=3)

次に、構築された学習器の予測性能を表 1 に示す。

また、テストデータ内で期間中の上昇率が高かった商品上位 3 件 (“Whim of Volrath” [14], “Pygmy Hippo” [15], “Flanking Troops” [16]) について、各学習器における予測結果と実際の価格履歴の推移を図 6、図 7、図 8 に示す。

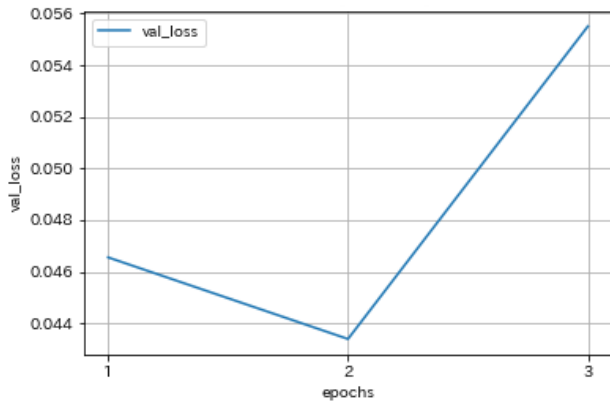


図 3 各ステップごとの val_loss(epoch 数=3)

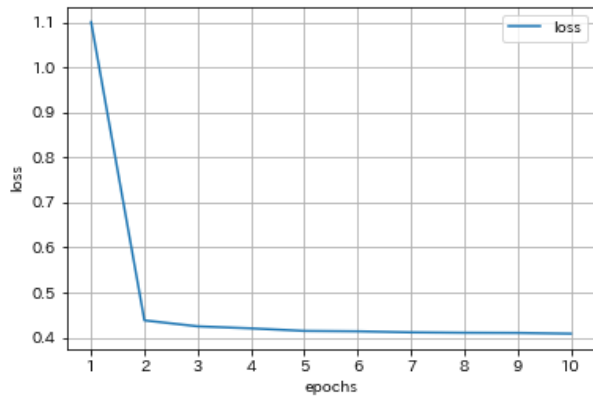


図 4 各ステップごとの loss(epoch 数=10)

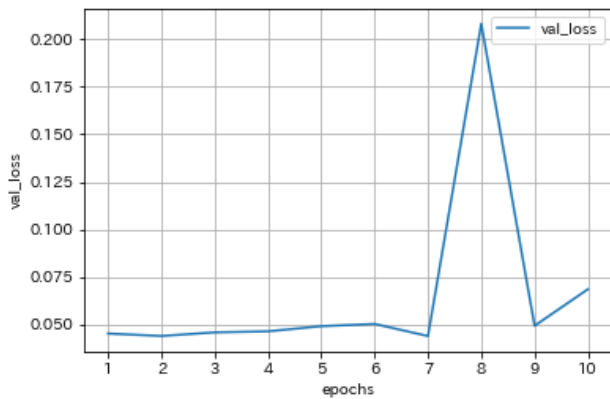


図 5 各ステップごとの val_loss(epoch 数=10)

表 1 各ステップごとの各評価指標の値

	epoch = 3	epoch = 10
MSE	17.32×10^{-2}	22.77×10^{-2}
RMSE	41.61×10^{-2}	47.72×10^{-2}
R^2	99.94×10^{-2}	99.92×10^{-2}

6 考 察

6.1 学習器の損失について

図 2, 図 4 について, どちらも 2 ステップ目で loss が減少し,

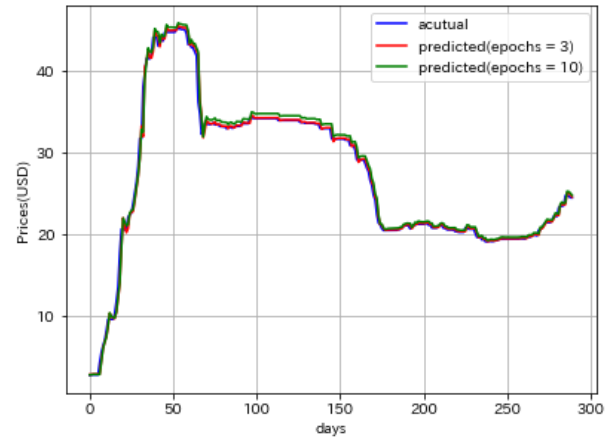


図 6 Whim of Volrath の予測結果

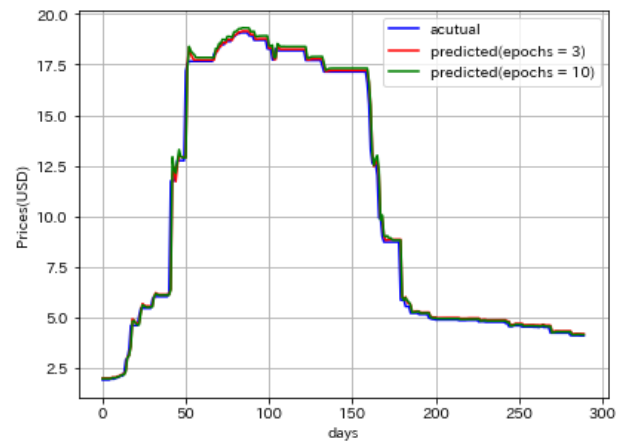


図 7 Pygmy Hippo の予測結果

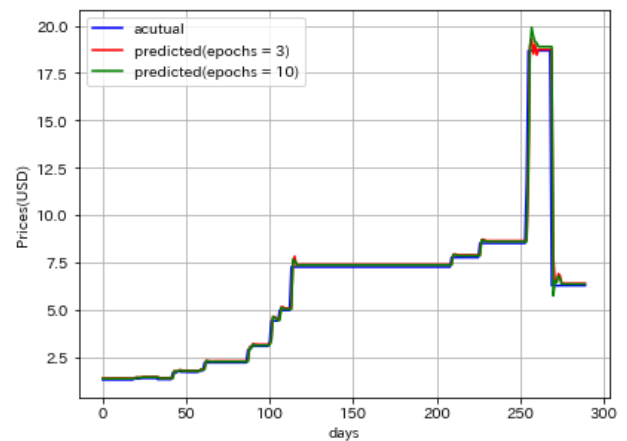


図 8 Flanking Troops の予測結果

その後緩やかに値が減少していることから, LSTM の学習に関して問題なく行えていることが示唆された. また, 図 3, 図 5 についても, val_loss も一部を除き 0.01 を超えない範囲の推移であることから, こちらも問題なく学習を行えているといえる. そのため, 本研究が目指していた TCG 中古販売価格への時系列分析の適用は, 提案手法を用いることで行うことができたといえる.

しかし、図 5 について、ステップ 8,10 において損失が大きくなっている。そのため、最適な epoch 数が 7 程度であり、それ以降の場合には過学習が起きている可能性がある。この結果から、epoch 数も含め今後パラメータチューニングに関する改善の余地があると考えられる。

6.2 学習器の予測性能について

表 1 について、1USD~150USD の価格範囲の中で平均二乗誤差、二乗平均平方根誤差が比較的小さな値となった点や、決定係数が 99×10^{-2} 程度となったことから、本研究の目指していた価格予測の精度として高い値を示し、提案手法の有効性を示すことができたといえる。これに関しては、図 6、図 7、図 8 から、それぞれの予測結果が実際の価格推移と乖離するものではなく、LSTM を用いた中古価格の時系列分析について有効であることが示唆された。

ただし、epoch 数が 3 の時と比べ 10 の場合の方が平均二乗誤差や二乗平均平方根誤差が大きいこと、テストデータに対する予測も epoch 数が 10 の場合実際の価格より一部高い区間がみられることから、精度については更なるチューニングの必要があるといえる。

また、一部パラメータが決め打ちである今回の実験において、チューニングがないにもかかわらず決定係数がかなり高いことから、過学習の可能性を含め、学習器の汎用性について検討する必要があることも示唆された。そのため、今後はその精度だけでなく汎用性の確認についても詳細に行い、予測性能を高めていく必要があると考えられる。

以上のことから、今回の目的であった LSTM による TCG 中古販売価格の時系列分析と予測については達成することができたといえる。しかし、その汎用性や精度について追及する余地が残る結果となった。だが、今回の提案について更なる展望を示すことができ、本研究の試みには意義があったといえる。

7 おわりに

7.1 ま と め

本研究では、TCG 市場における中古販売価格の時系列予測を目指し、LSTM を利用した学習と予測を行った。その結果、学習器の予測精度が高い値を示し、loss の推移やテストデータへの当てはまりからも、TCG 市場の中古販売価格について、時系列予測が行えることを示した。しかし、一部パラメータにおける損失の上昇や、予測精度がかなり高い結果となった点など、LSTM による分析には改善の余地があることも示された。だが、TCG に関する先行研究では扱われてこなかった時系列分析を取り上げて可能性を示すことができ、本研究の目的は達成することができた。

7.2 今後の展望

今後の展望について、まずパラメータチューニングによる精度向上、汎用性の確認が考えられる。パラメータチューニングについては、今後 Optuna [17] などを用いた機械的な探索を検討している。また汎用性の確認として、これ以降の日付の予測

や、対象外であった期間中に発売した商品など、更なる未知のデータについて予測し確認することを検討している。

また、今回は低価格帯や超高額帯に関する予測は行っていないため、それらに関しても同様のモデルで説明可能であるか考慮する必要がある。これらのデータに関しては、実際の例から決め打ちや一定の価格を維持し続ける傾向があるため、予測の際には価格帯ごとにモデルを切り替えるなどの方針を検討している。

加えて、今回の実験において取り扱っていない領域についても、予測に関して一考の余地があると考えられる。

今回は LSTM を用いて時系列分析を行ったが、LSTM 以外の時系列分析手法による予測の可能性も展望として挙げられる。予測精度の向上を目指し他の分析手法と比較することや、要因の分析を主眼置き、より解釈可能性が高いモデルによる分析も検討できる。

さらに Matthew らや穴戸らの研究では、価格履歴以外にも高騰要因がみられたため、それらの情報を追加した多変量時系列分析に関しても検討の余地があるといえる。

このように、TCG の中古販売価格の予測には課題が残るが、本研究で明らかにした内容を含め、今後も研究を継続し実験・考察していく予定である。

文 献

- [1] “令和 2 年度 産業経済研究委託事業（電子商取引に関する市場調査）報告書”，https://www.meti.go.jp/policy/it_policy/statistics/outlook/210730_new_hokokusho.pdf, (2022/02/14 アクセス).
- [2] “イーベイ・ジャパン，2021 年第 3 四半期の越境 EC トレンドを公開”，<https://www.ebay.co.jp/release/20211112/>, (2022/02/14 アクセス).
- [3] “マイコシンスの格子/Mycosynth Lattice”，<http://wonder.wisdom-guild.net/graph/Mycosynth+Lattice/?sym=median%2C1ow%2Ctrimmean&st=20190101&et=20190601>, (2022/02/14 アクセス).
- [4] “Magic: The Gathering”，<https://magic.wizards.com/en>, (2022/02/14 アクセス).
- [5] 穴戸建元，新美礼彦，“決定木学習に基づいた新製品発売時における中古価格高騰の要因分析”，2021-DBS-174, pp.1—6, 2021.
- [6] Matthew Pawlicki, Joseph Polin, Jesse Zhang, “Prediction of Price Increase for Magic: The Gathering Cards”, Stanford University CS229 Projects, pp.1—5, 2014.
- [7] 小松大起，黒岩丈介，小高知宏，諏訪いずみ，白井治彦，“電力使用量予測のための深層学習手法”，福井大学 学術研究院工学系部門研究報告，Vol. 70, pp.29—36, 2021.
- [8] 白澤元氣，加古真一郎，中村啓彦，“LSTM を用いた鹿児島市における集中豪雨イベントの再現”，AI・データサイエンス論文集，Vol. 2, No. J2, pp.893—901, 2021.
- [9] 宮崎邦洋，松尾豊，“深層学習を用いた株価予測の分析”，人工知能学会全国大会論文集，Vol. 31, pp.1—3, 2017.
- [10] 仲間拓哉，浦野昌一，“中長期的な経済指標予測のための LSTM によるエンコーダ・デコーダモデルの検討”，人工知能学会全国大会論文集，Vol. 34, pp.1—2, 2020.
- [11] “MTG JSON”，<https://mtgjson.com/>, (2022/02/14 アクセス).
- [12] “Card Kingdom”，<https://www.cardkingdom.com/>, (2022/02/14 アクセス).
- [13] “Keras”，<https://keras.io/>, (2022/02/14 アクセス).
- [14] “Whim of Volrath”，<https://gatherer.wizards.com/Pages/Card/Details.aspx?multiverseid=4742>, (2022/02/14 ア

クセス).

- [15] “Pygmy Hippo”, <https://gatherer.wizards.com/Pages/Card/Details.aspx?multiverseid=3740>, (2022/02/14 アクセス).
- [16] “Flanking Troops”, <https://gatherer.wizards.com/Pages/Card/Details.aspx?multiverseid=10486>, (2022/02/14 アクセス).
- [17] “Optuna”, <https://optuna.org/>, (2022/02/14 アクセス).