

トレーディングカードゲームの中古販売価格の要因分析

Analysis of Selling Price Factors for Used Trading Card Games

宍戸 建元[†] 新美 礼彦[†]

[†] 公立はこだて未来大学 情報アーキテクチャ学科 〒041-8655 北海道函館市亀田中野町 116 番地 2

E-mail: †{b1017154,niimi}@fun.ac.jp

あらまし 近年、トレーディングカードゲーム (以下、TCG) 市場は大きく広がりを見せている。TCG を形成する大きな要因のうち、ゲーム部分やコミュニティ以外の大きな要因として、中古カードの販売が存在する。この中古市場において、カードの販売価格に影響する要因は複数存在するが、要因それぞれがもたらす影響度合いが明確に分類されている例や、希少度を考慮した例は少ない。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、機械学習を利用した TCG の中古カードの価格形成要因の分析を行った。特に、カード自体が持つ属性を (1) 『カードゲームにおける性能属性』と (2) 『カードの希少度を表す属性』に該当するものに分類し、それぞれどのような属性が価格決定に寄与しているのか分析した。カテゴリ属性、数値属性を利用し、回帰分析と決定木 (CART) による分析の結果、性能属性よりも希少度属性の方が価格決定に寄与していることを示した。また、それぞれの属性の価格帯ごとの割合から、レアリティ、マナコスト、封入商品の発売年などの属性が価格に影響していることが示唆された。

キーワード データマイニング, 機械学習, 価格予測, EC サイト

1 背景と目的

1.1 背景

近年、トレーディングカードゲーム (以下、TCG) 市場は大きく広がりを見せている。本研究における TCG とは、「トレードなどによりカードの収集を行い、その中からルールに従って自由にカードを選び、カードを使用して遊ぶことのできるゲーム」のことを指す。

単にカードのコレクションを行うだけでなく、ゲームの要素を取り入れた TCG は、1993 年の「Magic: The Gathering」[1] から始まり、「遊戯王 OCG デュエルモンスターズ」や「カードファイト!! ヴァンガード」など、様々なタイトルが開発・販売されている。実際に日本において、玩具市場における TCG の売り上げは 2019 年で 1100 億円規模となっており、2018 年から 2019 年の間では成長率は 104.9% と、20 年以上たった今も成長を続けている大規模な市場である。[2]

TCG 市場を形成する要因のうち、ゲーム要素やコミュニティ以外のものとして、カードの中古販売が存在する。一般に TCG を遊ぶユーザは、ゲームで遊ぶために複数枚のカードを用意する必要がある。そのためユーザは、商品を購入する必要があるが、それらの商品の多くがランダムな確率でカードが封入されているため、本当に目当てのカードのみを手に入れた場合には、トレードや中古カード単品販売など別の手段を利用する必要がある。

TCG は特に中古カードの販売市場が発達し、「中古 TCG 専門店」としてカード販売のみを行う例もある。現在では、EC

サイト・アプリケーションを活用した電子商取引も活発になっており、中古 TCG 市場の内部でも情報技術が活用されている。

1.2 TCG 価格予測システムの構築

上述の中古販売市場の活発化から、本研究では TCG の中古カード自動値付けシステムの構築を最終目標としている。

中古 TCG 市場では、日々新しいカードが追加されるため、需要と供給が変化し、その価格も大きく変化する。また、カードゲームにおける性能は、実際にゲームについて深く理解しなければ判断が難しく、初心者や門外漢の人々にとって、是非を判断しながら価格をつけることは難しい。そのため、このシステムを実現することができれば、カードに関する知識が乏しい場合でも、カードの情報を入力することで簡便に値付けを行うことができる。また発展として、新たに発売されたカードと関連するような既存のカードを探し出し、それらの高騰予測までを扱い、市場の変化を予測するシステムも展望としている。

1.3 問題点

しかし、このシステム構築の前段階において、どのような要因が TCG カード価格に影響しているのか詳細に解明されていない、という問題がある。

TCG の中古販売市場では、カード 1 枚の販売価格は様々な要因によって決定されている。特定の年数を経た商品に高額な価格がつく事例や、イラストや言語の違い、カード光沢の有無などで価格が変動する事例がある。また、ゲーム上での性能が価格と直結する例も存在する。例えば、前週に 1 枚 30 円の最低価格で取引が行われていたカードが、相性の良い新しいカー

ドの発売により需要が増加し、次週では1枚2000円以上で取引される事例も存在する。このようにTCG市場では、価格に影響する要因が複数存在し、実際に価格が高騰した事例は存在する。しかし、要因それぞれが価格にもたらす影響が明確に分析された例は少ない。

1.4 目 的

前述の背景を踏まえ本研究では、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、機械学習を利用したTCG中古カードの価格要因分析を行う。

本研究では特に、その価格がカードの持つゲーム上の性能とそのカードの希少度によって説明できると仮定する。その上でカード自身の持つ属性を、(1)『カードゲームにおける性能属性』と、(2)『カードの希少度を表す属性』に分類し、それぞれの属性がどのように価格要因として作用するのかを、機械学習を用いて明らかにする。

2 関 連 研 究

商品の中古販売市場における価格要因分析の研究は、様々な先行研究が存在する。仁ノ平ら[3]の研究では、中古ファッションアイテムの価格予測を、ファッションブランドの出品履歴67,211件を対象とし、季節ごとの値下がり率を潜在クラスとして考慮して混合回帰モデルによる分析を行った。

また、越川ら[4]の研究では、中古車販売価格に対する既存研究を踏まえて、中古車の性能情報と価格情報を含んだデータセット157,611件を利用し、ランダムフォレストによる価格予測モデルを構築した。

表1は価格要因分析に関する先行研究と本研究との違いをまとめたものである。本研究の目指す価格予測システムの構築や価格要因分析は、他分野における価格要因分析の研究と同様の統計的手法による価格の分析から成る。しかし本研究は、TCGの中古市場を対象としたものであり、先行研究ではあまり扱われてこなかった商品を対象としている点で、独自性があるといえる。

表 1 価格要因分析に関する研究における本研究の位置づけ

論文	対象とする商品	データ数	研究内容
仁ノ平ら	ファッションブランドの出品履歴データ	67,211 件	潜在クラスを用いた混合回帰モデルによる分析
越川ら	中古車の出品履歴データセット	157,611 件	ランダムフォレストによる価格予測モデル構築
本研究	TCGの販売履歴データセット	41,236 件	線形回帰・決定木による価格要因分析

TCGの中古販売価格の分析・予測についても、いくつかの先行研究が存在する。

Loh[5]の研究では、「カードファイト!!ヴァンガード」のプレイヤー対象としたアンケート調査を行い、カードイラスト、ゲーム中の性能、カードの希少性、コミュニティの4つについて、中古販売市場における価格に影響を及ぼしていることを示した。本研究は、ここで示された「ゲーム中の性能」と「カードの希少性」について、カードの持つどのような属性が特に寄与しているのかを明らかにしようとする試みである。さらに、

この研究はアンケートインタビューである質的研究であり、本研究のデータや実価格から価格を分析する質的研究とは異なっている。

また、機械学習を利用したTCGの価格要因の分析には、Matthewら[6]の研究が存在する。この研究では、TCGの中古市場に見られる価格の高騰について、ロジスティクス回帰やSVMを利用し価格の変動を予測している。加えて、Dustinら[7]の研究においても、機械学習を利用したTCGの分析が行われている。この研究では、中古カード価格の予測において、カードのテキスト性能をn-gramsによる解析を行って利用し、カードゲーム上の性能から価格を予測している。これらの研究で扱われている予測は、本研究の目指している『カードゲームにおける性能属性』からの説明と合致している。しかし、『カードの希少度を表す属性』についてはあまり考慮されていないため、本研究では両方の属性について説明が可能であるか検証する。

このように、TCGの中古販売価格の分析はいくつかの参考研究が見られる。表2は、TCGの中古販売価格の先行研究についてまとめたものである。

表 2 TCGの価格要因分析に関する研究における本研究の位置づけ

	Loh ら	Matthew ら	Dustin ら	本研究
解析手法	アンケート調査	ロジスティクス回帰	n-gramsによるカードテキスト解析、回帰分析	線形回帰、決定木
対象データ	アンケート	2012年～2014年までのカードデータ、トーナメント使用カードデータ、価格データ	2015年までのカードデータ、価格データ	2020年11月9日までのカードデータ、価格データ
データ件数	107 件	13,608 件	14,352 件	カードデータ 41,236 件 価格データ 14,1434 件
対象タイトル	『カードファイト!!ヴァンガード』	『Magic: The Gathering』	『Magic: The Gathering』	『Magic: The Gathering』
研究内容	中古価格に関わる要因の分析	トーナメント利用履歴、価格履歴を用いた価格の変動要因を分析	テキストを含めた、カードの性能からの価格予測	カード価格の全体傾向と、価格に影響を及ぼす属性を明らかにする

この表からもわかるように、既存のTCGの中古販売の研究において対象にしているのはLohらのアンケート分析による質的研究か、Matthewら、Dustinらのような14,000件程度のデータによる量的研究である。しかし、本研究では「Magic: The Gathering」において、公式ゲームフォーマット上で利用できるすべてのカード41,236件を対象とし、既存研究を上回るデータ数における解析を行う。さらに、過去の量的研究とは異なり、カードの性能に関する属性のみではなく、カードの希少度も考慮に入れ分析を行う点が、既存研究と異なる部分である。また、その価格の要因についても価格の変動に特に注目するのではなく、カード価格全体の傾向について分析する点も異なる点である。

3 提 案 手 法

本研究では、Loh[5]の研究において示された中古販売市場への影響を与える要因について、特に「ゲーム中の性能」を「カードの希少性」について、カード性能を表す複数の属性から表現することが可能であると仮定し、分析を行う。

分析の際には、「Magic: The Gathering」で発行されたカードについて、その性能を複数の属性で記述したデータセットを用意する。その属性を(1)『カードゲームにおける性能属性』

と、(2)『カードの希少度を表す属性』に該当するものに分類し、それぞれどのような属性が価格決定に寄与しているのか、回帰分析と決定木学習によって明らかにする。

(1)『カードゲームにおける性能属性』とは、TCGのゲーム上の性能を決める要素である。例えば、カードタイプやカードテキスト、パワーなど、ゲームルールで参照されカードゲーム上での性能を表す属性を指す。

(2)『カードの希少度を表す属性』とは、TCGのゲーム上の性能にかかわらず、カードの希少性を決定する要素である。例えば、カードのレアリティや入手できる環境、再印刷の頻度など、そのカード自体の希少性の高さを表す指標となるものを指す。

また本研究では価格についても、『相場全体における価格要因の傾向』を分析するものとする。これは、Matthewら[6]の研究で確認している時系列的な価格変動とは違い、相場全体の傾向として高騰につながる要因を把握することが本研究の目的のためである。そのため、本研究における価格は、一定の日時において確認された価格を参照するものとし、その前後における価格の変動性については、本研究では扱わないものとする。

4 実験と考察

4.1 回帰分析による価格要因分析

4.1.1 実験概要

本実験では、(1)『カードゲームにおける性能属性』と(2)『カードの希少度を表す属性』について、カテゴリ属性・数値属性を利用し価格要因を説明が可能であるか、機械学習による回帰分析から評価を行った。

実験で使用したカードデータは、ゲーム上の公式フォーマットで使用不可能なカードや特殊なカードを除き、2020年9月29日時点までに価格データを取得することが可能であった41,236件を対象とした。これらのデータのうち、カードのもつ特性を表す属性は、“MTG JSON”[5]から「Magic: The Gathering」のすべてのカード情報が記述されたデータセットを利用し作成した。

作成したデータセットの属性のうち、ゲーム上で参照されると考えられる属性を公式ゲームルールと照らし合わせ考慮し、(1)『カードゲームにおける性能属性』と(2)『カードの希少度を表す属性』として分類し利用した。また、今回の実験で回帰分析に利用した属性は、そのうちのカテゴリ属性と数値属性であるもの24種類と価格データを利用した。表3、4に利用した属性と、図1に実際のカードとの対応の例を示す。

また、数値属性については“X”や“1+*”など、初期値がゲームの状況によって変化する旨が記載されている値が存在した。そのため、数値属性については可変であることを示す属性(例えば、power.Xなど)を新たに作成した。

また、価格データについても、同様に“MTG JSON”から中古小売価格データを取得し利用した。この価格データは、2020年6月28日から同年9月27日の期間において、海外のECサイトにおける販売価格141,434件を取得したものである。今

表3 実験に用いた(1)『カードゲームにおける性能属性』

属性	意味
colors	カード自身の色(赤:R, 白:W, 黒:B, 青:U, 緑:G)
textColor	カード自身の色以外にテキストで利用されている色(同上)
convertedManaCost	manaコストを色を考慮せず数値に変換したもの(0~16)
hasAlternativeDeckLimit	枚数制限の有無(4枚以上の使用が可能か否か)
layout	カードのレイアウト(normal, aftermath, split, flip, leveler, saga, transform, adventure, meldの9種)
loyalty	忠誠度(Planeswalkerタイプのみ使用, 0~7)
power	パワー(-1~16)
toughness	タフネス(-1~16)
types	カードのタイプ(Enchantment, Creature, Land, Instant, Sorcery, Artifact, Planeswalker, Tribalの8種)

表4 実験に用いた(2)『カードの希少度を表す属性』

属性	意味
hasContentWarning	人種差別を想起させるようなカードとして禁止されているか否か
hasFoil	フォイル版(特殊加工)の有無
hasNonFoil	ノンフォイル版(加工なし)の有無
isAlternative	特殊バージョンのカードの有無
isFullArt	フルアート版(特殊なイラスト)の有無
isPromo	プロモーション配布されたものか否か
isReprint	同名カードが再印刷されたか否か
isReserved	再録禁止リストにあるか否か
isStarter	スターターセット(商品)に入っているか否か
isTextless	テキストなしの特殊バージョンか否か
year	封入された商品の発売年度(1993~2020)
printings_num	同名カードの印刷回数(1回~174回)
borderColor	カードの枠線の種類(black, white, borderlessの3種)
frameVersion	カードレイアウトの種類(1993, 1997, 2003, 2015, futureの5種)
rarity	カードの商品内封入率(common, uncommon, rare, mythicの順に低い)

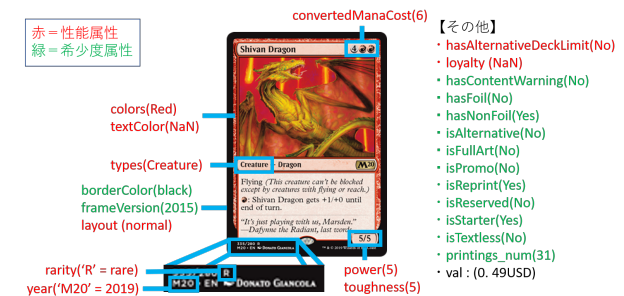


図1 実験に使用した属性と実際のカード[10]との対応の例

回は時系列的な傾向ではなく、相場全体における価格要因の傾向を確認するため、各カードの取得最新日の価格データのみに利用した。また、カードについても複数の店舗におけるカードの価格が存在したため、各店舗の最新価格取得日における価格同士の平均をカードの価格として利用した。その結果、価格は0.01USD~40084.505USDの範囲におけるカードを取り扱った。

データの前処理を行った後、データのうち9割を学習データ、1割をテストデータに分割し、学習データのうち価格(val)を目的変数、それ以外の属性を従属変数として、回帰分析を行った。本実験では、scikit-learn[9]に含まれているlinear_modelを利用し、線形回帰、ElasticNet、Ridge、Lassoの4つの回帰モデルによる学習器の作成を行った。

学習器の作成の際、(1)『カードゲームにおける性能属性』と(2)『カードの希少度を表す属性』の分類についてどのようにそれぞれが作用しているのか、(i)『カードゲームにおける性能属性』のみ、(ii)『カードの希少度を表す属性』のみ、(iii)両方の属性を利用した場合の3条件について、各4種類の回帰モデル

ルの作成を行った。

最終的に各回帰モデルの示した回帰係数を確認し、実際にどのような属性が価格要因として影響しているか確認を行った。また、(i)(ii)(iii)の条件下における回帰係数を確認し、性能・希少性のみを利用した場合と両方考慮した場合の違いについて確認を行った。

またそれらの学習器について、分割したテストデータを利用し、モデルの当てはまりを評価する決定係数 (R^2) とテストデータに対する二乗平均平方根誤差を求めることで、各属性から価格の説明が可能であるか否かを評価した。

4.1.2 実験結果

それぞれの回帰モデルの決定係数と二乗平均平方根誤差について、表 5, 6, 7 に示す。

表 5 (i)『カードゲームにおける性能属性』のみの各回帰モデルの性能

回帰モデル	決定係数 (R^2)	二乗平均平方根誤差
線形回帰	11.03×10^{-3}	158.2
ElasticNet	3.637×10^{-3}	158.8
Ridge	11.09×10^{-3}	158.2
Lasso	7.069×10^{-3}	158.5

表 6 (ii)『カードの希少度を表す属性』のみの各回帰モデルの性能

回帰モデル	決定係数 (R^2)	二乗平均平方根誤差
線形回帰	90.76×10^{-3}	151.7
ElasticNet	30.38×10^{-3}	156.4
Ridge	90.72×10^{-3}	151.7
Lasso	99.12×10^{-3}	151.0

表 7 (iii) 両方の属性を利用した各回帰モデルの性能

回帰モデル	決定係数 (R^2)	二乗平均平方根誤差
線形回帰	102.5×10^{-3}	150.7
ElasticNet	33.64×10^{-3}	156.4
Ridge	102.7×10^{-3}	150.7
Lasso	104.6×10^{-3}	150.5

また、性能属性・希少度属性・全属性の 3 条件において、Lasso 回帰において回帰係数を示した属性と、その回帰係数を表 8, 9, 10 に示す。なお、ここに示されている以外の属性については、属性削減を行う Lasso 回帰において回帰係数を示さず 0 だったものである。

表 8 Lasso 回帰において回帰係数を示した属性とその回帰係数 (性能属性のみ)

属性	線形回帰	ElasticNet	Ridge	Lasso
types_Artifact	5.115	6.924	5.424	40.11
convertedManaCost	-160.4	-1.384	-159.2	-10.39
types_Creature	-81.67	-5.239	-81.08	-10.91

表 9 Lasso 回帰において回帰係数を示した属性とその回帰係数 (希少度属性のみ)

属性	線形回帰	ElasticNet	Ridge	Lasso
isReprint	15.60	0	16.52	2.098
isReserved	460.2	16.40	461.5	414.5
frameVersion_1993	-4.960×10^{15}	14.82	41.00	53.22
rarity_rare	-2.020×10^{15}	9.833	5.979	8.558
frameVersion_1997	-4.960×10^{15}	-3.859	-26.26	-7.260

表 10 Lasso 回帰において回帰係数を示した属性とその回帰係数 (全属性)

属性	線形回帰	ElasticNet	Ridge	Lasso	属性分類
types_Artifact	5.337	6.372	5.471	28.03	性能属性
convertedManaCost	-171.5	-1.596	-169.8	-10.09	性能属性
types_Creature	-52.15	-4.852	-51.77	-7.928	性能属性
types_Enchantment	-50.19	-1.415	-49.82	-9.731	性能属性
isReprint	13.13	0	13.26	0.457	希少度属性
isReserved	456.6	16.36	455.2	412.1	希少度属性
frameVersion_1993	4.850×10^{14}	14.73	41.78	52.12	希少度属性
rarity_rare	-3.730×10^{14}	9.830	7.565	8.553	希少度属性
frameVersion_1997	4.850×10^{14}	-3.781	-23.31	-6.230	希少度属性

4.1.3 考察

a) 分類性能に関する考察

表 5, 6 より、(i)『カードゲームにおける性能属性』のみと(ii)『カードの希少度を表す属性』のみの各回帰の性能を比較すると、希少度のみを考慮した方が、決定係数がより高いため説明変数の当てはまりがよく、二乗平均平方根誤差もわずかではあるが抑えられていることから、TCG における価格要因としては、『カードゲームにおける性能属性』より『カードの希少度を表す属性』の方が強く影響していることがわかる。

しかし、表 7 について、両方の属性を考慮した場合においても、それぞれの回帰において、その決定係数は線形回帰・Ridge・Lasso において 100×10^{-3} 程度、ElasticNet では 30×10^{-3} とモデルへの当てはまりは低いものとなった。加えて、二乗平均平方根誤差はどのモデルにおいても 158 程度と、実際の価格帯が 0.01USD 単位であることから考慮すれば大きな誤差であり、予測精度は低いものとなった。このことから、性能属性と希少度属性の影響度合いを比較するという目標は達成できたが、数値属性・カテゴリ属性を用いて価格を説明が可能であるか、という点においてはうまく説明ができておらず失敗であるといえる。

失敗した理由について、一つは、線形回帰を利用した点そのものに問題があると考えられる。カード価格のデータ数上位 10 件の価格とそのデータ数を確認したところ、安い金額 (約 0.1USD) への極端な偏りが判明した。表 11 にその結果を示す。

このことから、目的変数である価格が均一に分布しておらず、正規化されていない状態では線形回帰によってうまく説明ができなかったと考察できる。また、(i)(ii)(iii) いずれの属性を用いた場合でも、価格帯に大きく偏りがあることから、線形回帰で表現できるモデルの複雑さではなく、実験に線形回帰を適用する手法自体が適切でないといえる。

b) 回帰係数に関する考察

前述した分類性能に関する考察から、希少度属性の方がその

表 11 価格データ数上位 10 件の価格とそのデータ数

カード価格 (USD)	データ数
0.101	538
0.099	501
0.096	464
0.094	455
0.104	451
0.091	407
0.106	391
0.125	374
0.109	355
0.089	339

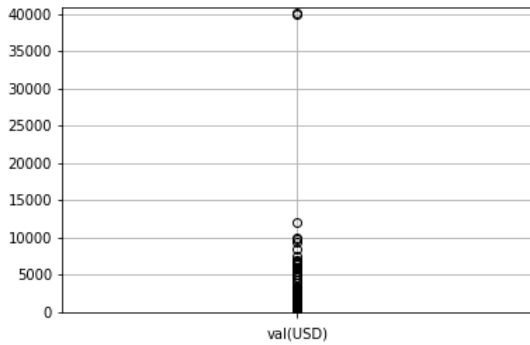


図 2 販売価格の箱ひげ図 (0.01USD~40084.505USD)

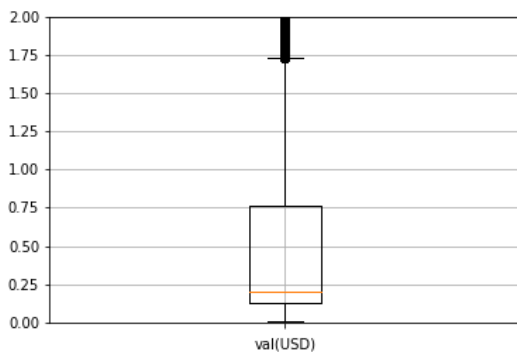


図 3 販売価格の箱ひげ図 (0.01USD~2.00USD)

価格への影響度が高いことを示せたが、予測精度が乏しいため、求められた属性ごとの回帰係数についてもその妥当性の判断が難しく、どの要因が特に寄与しているのかを定量的に判断することは難しい。

しかし、分析の結果いくつかの属性について、属性選択を行う Lasso 回帰においても回帰係数が見られた。性能属性では types_Artifact が正の回帰係数、convertedManaCost、types_Creature が負の回帰係数を示した。また、希少度属性では、isReprint、isReserved、frameVersion_1993、rarity_rare が正の回帰係数を示し、frameVersion_1997 のみが負の回帰係数を示した。これらの属性の妥当性については、後述する実験 4.3 においてその価格帯ごとの傾向を比較して判断を行った。また、全属性を考慮した場合にのみ、性能属性のみで作成した分析では回帰係数を示さなかった types_Enchantment が負の回帰係数を示し、興味深い結果となった。

4.2 決定木学習 (CART) による分類実験

4.2.1 実験概要

実験 4.1 の価格要因分析の結果から、回帰分析以外の手法を用いて検証するため、本実験では (1)『カードゲームにおける性能属性』と (2)『カードの希少度を表す属性』について、決定木学習を用いて各属性による説明が可能であるか分析を行った。

実験で使用したデータは、4.1 の実験と同様に、前述したカードデータセット 41,236 件とその価格データ 141,434 件を対象とした。使用したデータセットの属性は前述した表 3 と表 4 と同様のものを利用した。その際、数値属性のみの前処理段階での 01 正規化をせずに決定木を構築した。

また、クラス分類問題として決定木を適用するため、価格のうち第 3 四分位点であった 0.765USD を基準とし、それより高価なカードを “High” クラス、それ以外を “Low” クラスとする 2 クラス分類として決定木を構築した。

本実験では、決定木の構築に scikit-learn に含まれている Decision Tree Classifier(CART) を用いた。その際には、分類の基準にエントロピーを利用し、それ以外のいくつかのパラメータについてはデフォルトの値を用いた。この実験においても、前述の実験 4.1 のように、(i)『カードゲームにおける性能属性』のみ、(ii)『カードの希少度を表す属性』のみ、(iii) 両方の属性を利用した場合の 3 種類の学習モデルの作成を行いその比較を試みた。

最終的に、実験に使用したデータセットのうち、9 割を学習データ、1 割をテストデータとし、学習データを用いて決定木を構築した。また、前述した決定木の評価指標について、陽性を “High” クラスとして計算し、テストデータを利用した予測結果の評価を行った。その結果より、構築した決定木の性能からそれぞれの属性による説明が可能か否かを評価した。

4.2.2 結果

構築された決定木の予測結果の評価について表 12 のように示す。

表 12 各条件に置ける決定木の予測結果の評価尺度とそのスコア

評価尺度	(i) 性能のみ	(ii) 希少度のみ	(iii) 全属性
Accuracy	0.741	0.806	0.806
Precision	1.000	1.000	1.000
Recall	0.022	0.331	0.331
F 値	0.043	0.470	0.470

また、全属性を利用して構築した決定木の 3 段目までを図 4 に示す。

4.2.3 考察

表 12 の各条件における決定木の分類性能評価について、性能属性の方が Accuracy が大きく、Recall の値も 0.022 と極端に低いことから、性能属性よりも希少度属性の方が全体の価格要因として寄与していることが示された。この結果は、図 4 の全属性を用いて構築した決定木のうち 3 層目までのすべての属性が希少度属性であることから、希少度属性が特に影響度が高いことを示唆している。

また、構築された決定木を確認したところ、分類ルールとして

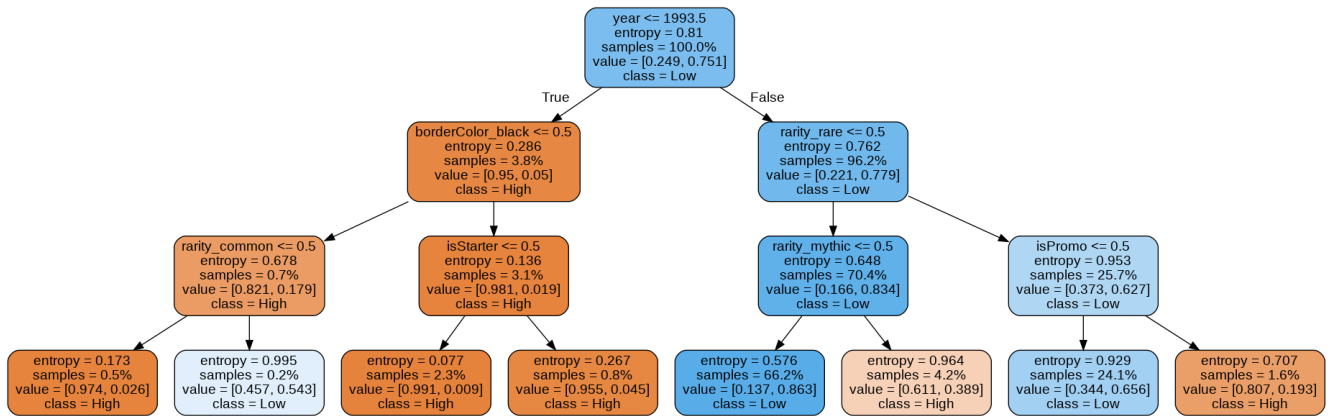


図 4 全属性を利用して構築した決定木 (3 段階目まで)

用いられた属性として year や borderColor.Black, isStarter, rarity_common, rarity_mythic などの、回帰分析の実験において示されなかった属性が分類に用いられており、興味深い結果となった。これらについても、後述する実験 4.3 において価格帯ごとの傾向を確認しながら評価を行った。

しかし、最終的に “High” に分類されたサンプルを確認すると 9%程度しか残らなかった。これについて、前述したように価格の極端な偏りも原因であると考えられる。だが、決定木の分類性能について、F 値の値がどの決定木においても 0.47 を上回らず決定木の分類性能が低いことから、数値属性・カテゴリ属性を用いて価格を説明することが可能であるか、という点においてはうまく説明ができておらず失敗であるといえる。また、分類性能が乏しいものとなったのに対し Precision が 1 となったことから、ほとんどのテストデータについて “Low” に分類される決定木になったといえる。これらのことから、分類に使われている属性それぞれについて、この結果から影響度合いを十分に比較することは難しいといえる。

4.3 価格帯ごとの属性傾向の確認実験

4.3.1 実験概要

実験 4.1, 4.2 より、回帰や分類に利用されてきた属性が確認できたが、どちらも精度が乏しいものであり、それらにおける属性の妥当性を定量的に検討することは難しいことが示された。そこで、本実験では実験 4.1, 4.2 で示された属性について、価格帯ごとのデータ割合を調べ、実際のデータ中において、その属性に関する傾向がみられるかどうか実験を行った。

実験で使用したデータは、実験 4.1 や 4.2 と同様に、前述したカードデータセット 41,236 件とその価格データ 141,434 件を対象とした。使用したデータセットの属性は前述した表 3 と表 4 と同様のものを利用した。

今回は、それらのデータについて、価格 (val) の値ごとに 4 つの価格帯ラベルに分類した。表 13 は価格帯とその分類クラス、該当するデータ数を表す。各価格帯の分類は、実際の相場価格を参考にラベル境界を決定した。実験では、このラベルごとに確認した属性における各値の割合を確認し、その妥当性を検討した。また、全体における割合についても同時に確認した。

表 13 価格 (val) の値と該当する価格帯ラベルとの対応

価格帯ラベル	価格 (val) の範囲	データ数
Low	0.01 ~ 0.74USD	30920
Middle	0.75 ~ 10.0USD	8178
High	10.01 ~ 150.0USD	1796
More.High	150.01USD ~ 40084.505USD	342

また確認を行った属性として、実験 4.1 おいて Lasso 回帰が回帰係数を示した属性である、types_Creature, types_Enchantment, types_Artifact, isReprint, isReserved, convertedManaCost, frameVersion_1993, frameVersion_1997, rarity_rare の 9 属性を利用した。また、実験 4.2 の決定木において、前述した 9 属性以外に分類に利用された、borderColor.Black, year, rarity_common, rarity_mythic, isStarter, isPromo の 6 属性についても確認を行った。

最終的に、以上の属性について価格帯ごとのユニークな値の割合を調べ、実際に実験 4.1, 実験 4.2 で示されたような属性に、価格との関係性があるのか確認した。

4.3.2 結果

属性の値と価格ごとの割合について、特に High, More.High と Low, Middle について異なる傾向を示したものを、表 14, 15, 16, 17 に示す。カテゴリ属性の割合を表した表 14, 16 ではそれぞれの価格帯における True, False の割合を表す。数値属性の割合を表した表 15, 17 では割合の高い上位 5 件について、その値の割合を表す。

表 14 rarity_rare におけるデータと価格帯ごとの割合

	ALL	Low	Middle	High	More.High
False	72.78%	82.21%	45.30%	47.32%	11.40%
True	27.21%	17.78%	54.69%	52.67%	88.59%

4.3.3 考察

表 14 について、rarity_rare=True の割合が Low の段階では 17.78%と少数であったが、Middle, High, More.High では True の方が多い結果となった。さらに、More.High では 88.59% が True を示す結果となった。このことから、rarity_rare, つまりカードのレアリティがレアに属するものについて価格との

表 15 convertedManaCost(0~16) におけるデータ上位 5 つと価格帯ごとの割合

ALL		Low		Middle		High		More.High	
値	割合	値	割合	値	割合	値	割合	値	割合
3	21.30%	3	22.06%	3	19.61%	3	17.53%	0	29.23%
2	19.41%	2	20.52%	4	16.39%	0	15.86%	2	20.17%
4	16.50%	4	16.65%	2	16.05%	2	15.59%	1	15.20%
1	11.40%	1	11.30%	0	12.37%	1	15.08%	4	12.86%
0	11.28%	0	10.53%	5	11.21%	4	15.03%	3	12.86%

表 16 frameVersion.1993 におけるデータと価格帯ごとの割合

	ALL		Low		Middle		High		More.High	
False	89.98%		94.02%		85.14%		58.74%		4.67%	
True	10.01%		5.97%		14.85%		41.25%		95.32%	

表 17 year(1993~2020) におけるデータ上位 5 つと価格帯ごとの割合

ALL		Low		Middle		High		More.High	
値	割合	値	割合	値	割合	値	割合	値	割合
2019	9.77%	2019	10.37%	2019	8.96%	1993	30.40%	1993	85.96%
2018	7.05%	2018	6.97%	2020	8.35%	2020	9.57%	1994	9.35%
2020	6.29%	2017	6.39%	1993	8.31%	1994	9.24%	1999	2.92%
2017	6.22%	2020	5.62%	2018	7.71%	2018	6.62%	1998	1.16%
2016	5.24%	2016	5.35%	2017	6.45%	1999	5.40%	2020	0.29%

関係性が高い属性であるといえる。また rarity_rare は、実験 4.1 において回帰係数が正の値を示したことから、レアリティがレアであればより価格が高くなるといえる。これは、実際の商品においても、レアリティという指標は商品内でのカードの封入率を表し、レアリティがレアのカードは封入率が低いものである。そのため、レアリティはそのカード自体の供給量を表す指標と考えることも可能であり、カードの価格との影響があることには妥当性があるといえる。しかし、回帰分析の際にはその回帰係数の大きさは特出したものではなかったが、実際には ALL において少数属性であり、価格ごとの割合をみた際に大きな差がみられ、興味深い結果となった。

表 15 について、convertedManaCost の各価格における上位 5 つは、All や Low, Middle の際には、3 が上位を占めていたが、More.High では 0 が 30%近くを占めていた。このことから、最上位価格層については特に軽量の convertedManaCost、つまりマナコストが軽いものがより重要視されると示された。また convertedManaCost は、実験 4.1 において回帰分析の際に負の相関係数を示したが、これは実際に最上位価格層では小さい値が重要視されており、軽ければ軽いほど価格が上がり、負の回帰係数を示した原因となったといえる。これについては、実際のゲーム上の性能においても、マナコストの値が小さいほどゲーム中で利用しやすい傾向がある。そのため、性能による需要としてマナコストが小さいほどカード価格が高くなることには妥当性があるといえる。

表 16 について、frameVersion.1993 の各価格における割合は、Low において True が 5.97%、More.High において True が 95.32%となった。また、表 17 について、各カードが封入された商品の発売年を表す year の割合上位 5 件が、High, More.High において 1993 年が一番高い結果となった。このことから、価格の要因として発行年度は価格との強い関係があるといえる。実際に、Low や Middle の場合では、最上位の 2019 でも 11%を超え

ないのに対し、1993 年について High では 30.40%、More.High では 85.96%となっていることから、特に 1993 年が価格に対して関係しているといえる。これは実験 4.2 の全属性を利用して構築した決定木において、根ノードの分岐条件が $year \leq 1993.5$ 、つまり 1993 年か否かを確認していたことから、実際の分析結果と一致するものとなっている。さらに、1993 年は「Magic: The Gathering」が発売された最初の年であるため、特に最初期に出たカードについて中古販売価格が高くなることが示唆された。

以上のことから、線形回帰や決定木で確認できた属性の一部について、実際の割合から妥当性があることが示唆され、実験は成功したといえる。

5 まとめ

本研究では、中古 TCG の価格要因分析を行うため、カード自体が持つ属性を (1)『カードゲームにおける性能属性』と (2)『カードの希少度を表す属性』に該当するものに分類し、それぞれのような属性が価格決定に寄与しているのか分析を行った。

回帰分析と決定木による分析を通して、性能属性よりも希少度属性の方が価格決定に寄与していることを示した。また、価格帯ごとに属性の持つ値とその割合を確認し、レアリティ、マナコスト、封入商品の発売年が価格に影響していることが示唆された。しかし価格予測の点においては、価格の分布が 0.1USD 付近に極端に偏っていることや、回帰分析の結果を用いた価格予測の精度や分類木の性能が著しい結果となったことから、属性同士の影響を比較し精査することは達成できなかったものといえる。また、本研究の目指す TCG の価格予測システムについても、その予測精度から数値予測ができていたとはいえず、より複雑なモデルでなければ TCG の予測が困難であることが判明した。

研究の展望について、実験で示された内容から性能属性に関する検討や、別の機械学習手法による分析などがあげられる。性能属性について、今回の実験では価格にそれほど寄与していない結果となったが、実例として性能による高騰が起きた例が確認されていることから、性能属性と価格との影響について再検討する必要がある。また実験中においても、types.Enchantment のような性能と希少性を同時に利用することで確認された属性も存在し、属性の影響について完全に分析できたとは言えないため、別の手法による分析を通してより解明を進める必要があるといえる。

また、本研究では扱っていない部分についても、展望として一考の余地がある。その例として、今回取り扱っていないテキスト属性を扱った価格要因分析や、時系列的な価格変動の要因分析などがあげられる。特に性能属性は、その評価がトーナメントでの利用率・ゲームルールの変更などによっても変化するため、それぞれの要因の価格への影響を検証することも課題である。

このように、TCG の中古価格要因の解明には多くの課題が残るが、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・検証を行う予定である。

文 献

- [1] 「Magic: The Gathering」, <https://mtg-jp.com/>, (2021/02/11 アクセス).
- [2] “日本玩具協会 2019 年度国内玩具市場規模”, https://www.toys.or.jp/pdf/2020/2019_data_zenpan.pdf, (2021/02/11 アクセス).
- [3] 仁ノ平 将人, 三川 健太, 後藤 正幸, “販売履歴データに基づく中古ファッションアイテムの販売価格予測モデルに関する一考察”, 情報処理学会論文誌, Vol.60, No.4, pp.1151-1161, 2019.
- [4] 越川 翼, 林 高樹, “中古自動車の価格要因分析”, 慶應義塾大学大学院経営管理研究科修士課程学位論文, 2018.
- [5] Loh Zhen Yu, “Factors Affecting Prices Of Singles in Card-fight !! Vanguard”, PhD thesis, UTAR, 2017.
- [6] Matthew Pawlicki, Joseph Polin, Jesse Zhang, “Prediction of Price Increase for Magic: The Gathering Cards”, Stanford University CS229 Projects, 2014.
- [7] Dustin Fink, Benjamin Pastel, Neil Sapra, “Predicting the strength of Magic: The Gathering cards from card mechanics”, Stanford University CS229 Projects, 2015.
- [8] “MTG JSON”, <https://mtgjson.com/>, (2020/02/11 アクセス).
- [9] “scikit-learn”, <https://scikit-learn.org/stable/>, (2021/02/11 アクセス).
- [10] “Shivan Dragon”, <https://gatherer.wizards.com/Pages/Card/Details.aspx?multiverseid=469888>, (2021/02/11 アクセス).