

# 機械学習を利用した創作物の知的財産的価値への寄与の評価のための情報理論的な測度について

金子 格<sup>†1</sup>

**概要:** 機械学習の進歩により、機械学習を利用して創作された絵画、音楽、小説などの「創作物」(Machine aided creative works- MACW) を利用できる可能性が高まってきた。機械学習を利用した「創作物」の生成手法とその品質は、現在急速に進歩しており、近い将来幅広い範囲で高品質な作品の利用が可能になると予想される。本論文では、このような機械学習を利用して創作された創作物の知的財産的価値への寄与を評価する場合における諸問題を論ずる。まずこれらの知的財産的価値への寄与の評価において、入力(操作やデータ)の情報量を基礎として、なんらかの測度が与えられないかを検討する。検討の結果として筆者は、入力の単純な「情報量」による測度はMACWの各入力を持つMACWの知的財産的価値への寄与を必ずしも適切に反映しないと結論する。次に、生成されたMACWの人による評価をMACWの入力の知的財産的価値への寄与の評価に反映する方法を検討する。人による評価による創作物の価値は、MACWの入力の多くの要因に分解され得る。そこで、統合されたMACWへの人による評価に対する複数の要因の寄与を評価に組み入れるための簡単な数量モデルを提案する。

**キーワード:** 機械学習, 創作, 知的財産, アルゴリズム情報量, 市場評価

## An information theoretical measure for the evaluation of contribution to the value of intellectual properties of works produced using the Machine Learning Techniques

ITARU KANEKO<sup>†1</sup>

**Abstract:** Thanks to the advancement of machine learning technologies, the possibility is rising for the use of machine-aided creative works (MACW), such as paintings, music and literature, generated using machine learning.

The production method of "works" by machine learning and its quality are rapidly progressing now, and wide range of applications with superior quality are expected to expand. In this paper, we will discuss some aspects of the evaluation of MACW as a kind of works, especially on how to evaluate the complex contribution to the value associate with intellectual properties. At first, we first discuss whether any measure can be given as the contribution, based on the amount of information of input, for the evaluation of the value associate with these intellectual properties. As a result of the discussion, the author concludes that these contributions do not exactly reflect values relates to the intellectual property evaluated by human.

Then, we will consider how to use value of generated MACW evaluated by human for the evaluation of contributions to the value associated with MACW. The human evaluated value of final content may be decomposed to the many factors of MACW. Furthermore, we propose the numerical evaluation model to evaluate the contribution of multiple elements to the human valuation of integrated MACW.

**Keywords:** Machine Learning, Creation, Intellectual Property, Algorithmic Information, Market value

### 1. はじめに

機械学習の進歩により、これまで人間が自らつくり出してきた絵画、音楽に変わり、機械学習を利用した「創作物」絵画、音楽、小説などを広く利用する可能性が高まってきた。

機械学習を利用した「創作物」の一例を図1に示す。この例はAlexander MordvintsevのDeepDream[1]により生成された画像でありWikimedia Commonsに収録されているものを引用した[2]。DeepDreamは畳み込みニューラルネットワークを用いて画像のパターンや特徴を分析再合成するコ

ンピュータビジョンプログラムである。Python, Anaconda, Café上で実現されたソースコードがGitHubで公開されている[3]。

コンピュータによる創作の可能性は古くから議論されてきた[4]。現状のDeepDreamは絵画表現として興味深きはあるが、ごく限られた特殊な画像しか生成することはできない。しかしこれは端緒にすぎず、機械学習を利用した「創作物」の生成手法は急速に発達し、生成できる作品は質量ともに急速な進歩を続けている。近い将来、これまでは人間がつくりだしてきた絵画、音楽、小説に代わり、幅広い分野での利用が広がると予想される。その知的財産的価値

<sup>†1</sup> 名古屋市立大学  
Nagoya City University.

への寄与の評価は、当然解決すべき課題だろう。

以下では機械学習を利用した「創作物」を Machine-aided creative content, MACW と呼ぶこととし、その知的財産的価値への寄与を評価する場合の諸問題を議論する。

ここで知的財産的価値への寄与とは、MACW の創作に際して関与したすべての入力がある MACW の市場的「価値」にどれだけ寄与したか、ということの意味する。そもそもそのような寄与の確立した定義はなく、法的な知的財産権は必ずしも知的財産的価値への寄与と対応しない。本論文は法的な知的財産権について論ずるものではなく、コンテンツ産業振興の手段として知的財産的価値への寄与に何等かの指標を与えるか、という問題を論ずるものである。

MACW の生成方法は様々である。生成には学習に用いた「参照情報」、「学習アルゴリズム」、学習によって蓄積された「蓄積情報」機械学習を利用して創作する者の「指令」や「入力データ」など様々な情報源が寄与し、その寄与の割合も様々である。

細部が参照したデータの単純コピーに近い場合もあるし、学習の結果得られた統計的な特徴のみから生成する場合もある。また、学習アルゴリズムや生成アルゴリズムは、単純な数値処理である場合もあるし、それ自体が質の高い出力を生む能力を持つ高度なアルゴリズムである場合もある。「指示」が処理の開始命令のみの場合もあれば、絵筆の細部やバランスを細かく指定する場合もある。MACW の知的財産的価値への寄与の評価では、これらを寄与の有無という二値的な扱いではなく、寄与の多寡を定量的に評価すべきではないだろうか。その場合、どのような定量化が可能だろうか。

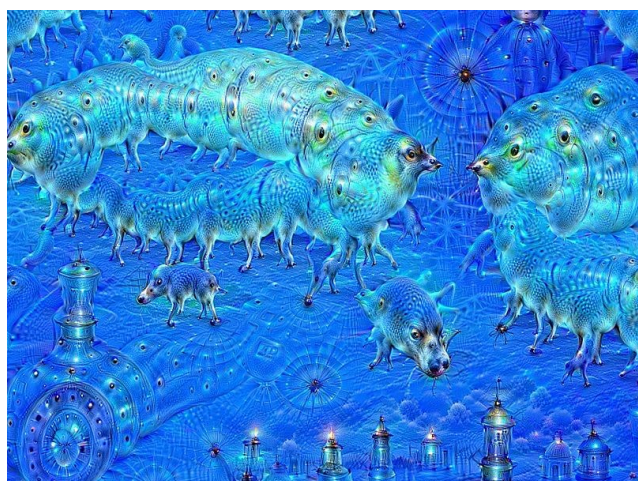


図 1 機械学習による作画例(Aurelia-auria-3-0009.jpg,[2])

Figure 1 The example of image created machine learning(Aurelia-auria-3-0009.jpg,[2]).

本論文の前半では、MACW への入力の寄与の評価において、入力される情報量を基準とすることができないかを、まず論ずる。筆者は、単純な「情報量」による測度は人間

による創作物の知的財産的価値への寄与に、常に適切な評価を与えるものではないと結論する。

後半では、単純な情報量による評価の問題点を解決するために、アルゴリズム情報量と、人による評価を反映する方法を提案する。

## 2. 機械学習による創作物生成のモデル

### 2.1 機械学習による創作物生成の参照モデル

筆者は人工知能(AI)や機械学習を利用した「創作物」の知的財産的価値への寄与の分析に共通の参照モデルを用いることを提唱している[5][6][7]。ここで参照モデルとは、様々な相似なシステムの仕組みや特性を比較するために用いる共通の枠組みを意味する。参照モデルにあてはめることで、各構成要素の定義を明確に与え、議論を厳密なものとすることができる。

本論文では、MACW の創作も同じ構造のモデルで議論できると考え、図 2 に示すモデルを参照モデルとして使用する。ただし、上記参考文献[5][6][7]では出力は「知財化候補データ」としていたが、本論文では出力は MACW に変更しその定義も改めた。

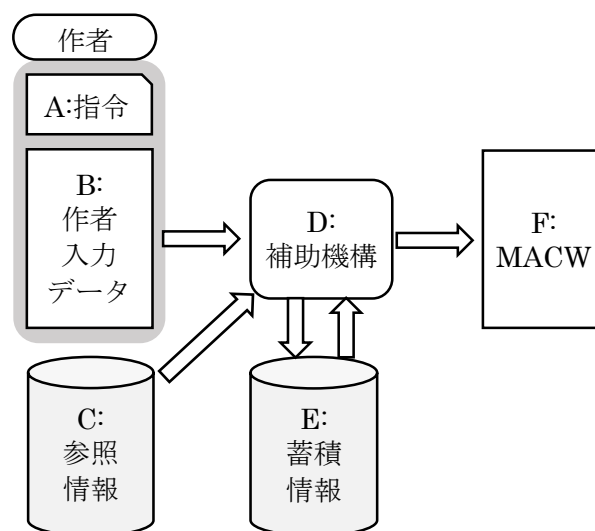


図 2 機械学習を利用した創作の参照モデル

Figure 2 The reference model of creative work using machine learning.

モデルの各構成要素の定義を説明する。F:MACW は出力、すなわち機械学習を利用した「創作物」である。人間の作者は A:指令と、B:作者入力情報を入力することにより、D:補助機構を操作して創作を行う。C:参照情報とは補助機構が創作物の生成時に外部から参照する情報である。E:蓄積情報とは機械学習の場合には、学習結果、たとえば機械学習 P のネットワークの係数である。

極端な例として従来の人間による創作活動も本モデルにあてはめることが可能であり、それは C, D, E の寄与がゼロである場合とみなせる。その場合、作者はたとえば画像

ならば描画操作や色指定など、文章ならばテキストそのものを B: 作者入力データとして入力する。

その反対の極端な例として DeepDream を参照モデルにあてはめた場合には作者(DeepDream を操作する人)は A: 指示や B: 作者入力データをほとんど与えない。最小限の A: 指示を与えることで、機械学習によって学習した学習結果 D: 蓄積情報と、C: 参照情報を利用して F: MACW の画像を生成する。なお法的には DeepDream による MACW の「作者」は「操作者」であって「作者」とは呼べないという主張があるが、本論文では「作者」をモデル上の要素としてのみ扱い法的判断は論じないこととする。

これら両極端の場合の知的財産的価値への寄与は比較的単純である。その中間的な場合はより複雑になるだろう。一般的には機械学習が部分的に利用され、多くの寄与が様々な比率で関与する。

### 3. 単純な情報量による評価

ここではまず、MACW への寄与を単純な情報量により評価することが可能であるかについて検討する。

#### 3.1 一様分布の蓄積情報

議論を容易にするために、画像ではなく文字列の生成を考える。図 3 にあらためて参照情報と蓄積情報のみから出力が得られている状況を示す。文字列の生成を行うものとし、均等確率の  $E_1$ : 蓄積情報と  $C_1$ : 参照情報から  $F_1$ : MACW を生成するとする。均等確率の蓄積情報  $E_1$  を表 1、参照情報:  $C_1$  を表 2 に示す。蓄積情報:  $E_1$  には、様々な長さの文字列が集積され各  $s$  と  $l$  について対応する文字列が蓄積情報として蓄積されている。文字列の分布はすべての文字列パターンで均等である。 $C_1$ : 参照情報はその蓄積情報をピックアップする符号である。適切な符号により任意の長さの任意の文字列を選択することが可能で、これにより任意の英文を生成することができる。この場合  $E_1$  は単純に  $C_1 \rightarrow F_1$  の写像を与えているだけである。したがって  $F_1$ : MACW の知的財産的価値に  $E_1$  は寄与しておらず  $C_1$  だけが MACW に知的財産的価値に寄与していると考えられる。

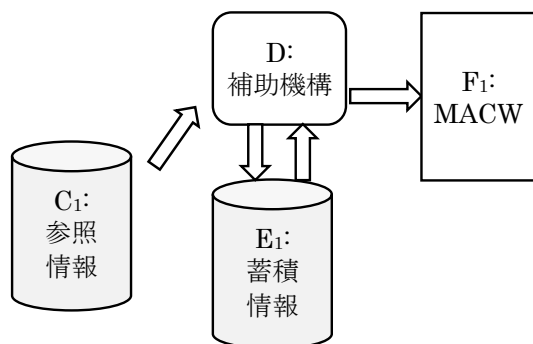


図 3  $C_1, E_1$  から  $F_1$  を生成するモデル

Figure 3 A model using stored information  $E_1$

画像の場合にも同様な状況が考えられる。8x8 の画素の単位で大量の画像の画素間の相互相関を学習させると  $E_1$  には DCT に近い基底(正確には KL 変換)が生成されることが知られている。DCT 基底は画像の特徴を抽出してはいるが、DCT 基底を使ってあらゆる画像は表現可能である。DCT 基底は周波数から空間への写像を与えているだけで MACW の知的財産的価値に寄与しておらず  $C_1$  が  $F_1$ : MACW の知的財産的価値に寄与していると考えられる。

表 1 一様分布を持つ蓄積情報  $E_1$

Table 1 Stored information  $E_1$  with uniform distribution

	$l=1$	$l=2$	$l=3$	$l=4$
$H_1[\text{bit}]$	4.75	9.5	14.25	19
$s$				
1	A	AA	AAA	AAAA
2	B	AB	AAB	AAAB
...	...	...	...	...
26	Z	AZ	AAZ	AAAZ
27		BA	ABA	AABA
28		BB	ABB	AABB
...		...	...	...

表 2 参照データ  $C_1$  と符号長

Table 2 Reference data  $C_1$  and code length.

Command	$(s, l)$	出力	$H_2[\text{bit}]$
1	1, 1	I	4.75
2	2, 12	AM	9.5
3	2, 374	OK	9.5

#### 3.2 非一様分布の蓄積情報

次に図 4 に非一様分布の蓄積情報  $E_2$  を利用するモデルを示す。蓄積情報  $E_2$  は表 3 に示す非一様分布英文によって構成されている。蓄積情報  $E_2$  には英文が蓄積されているがその出現順序は英文中で利用される頻度の順に並べられている。したがって有名なフレーズほど上位になる。さらに確率分布が不均等であることを利用して可変長符号化を行っている。頻度が多い文に短い可変長符号を付与し、その bit 数を  $H_1$  に示している。

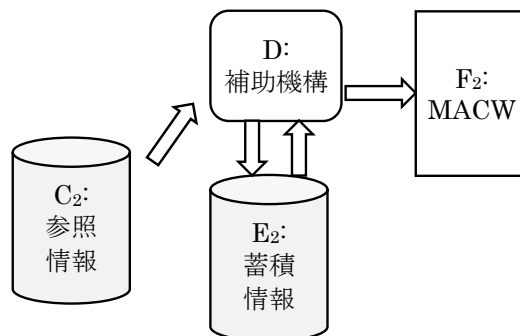


図 4 非一様分布の蓄積情報  $E_2$  から  $F_2$  を生成するモデル

Figure 4 A model which generates  $F_2$  using stored data  $E_2$  with nonuniform distribution

表 4 に参照データ  $C_2$  と符号長を示す. Command とそれにより示す指示により出力文字列を示し, また出力の bit 数として空白を含む文字数 $\times 4.75\text{bit}$  を  $H_2$  として示している. Command は均等確率の場合と同様に, 蓄積情報の選択を指定する. ここでは 2 つの文章を選んで, そしてこの 2 つの文章を接続すれば Simon & Garfunkel, “The Sound Of Silence” の歌詞の冒頭部分になる.

表 3 非一様分布を持つ蓄積情報  $E_2$ Table 3 Stored information  $E_2$  with non-uniform distribution.

$s$	$H_1[\text{bit}]$	Sentence
1	1	HELLO
2	1.5	YES
3	2	NO
4	2.3	I LOVE YOU
5		WHO ARE YOU
...		
2145656	21	HELLO DARKNESS MY OLD FRIEND
...		
369310405	28	IVE COME TO TALK WITH YOU AGAIN
...		

表 4 参照データ  $C_2$  と符号長Table 4 Reference data  $C_2$  and code length.

Command	出力	$H_2[\text{bit}]$
2145656	HELLO DARKNESS MY OLD FRIEND	137.8
369310405	IVE COME TO TALK WITH YOU AGAIN	152.1
...	...	

この場合  $E_2$  の上位には単純な文が多く収容されているが, やがて大量の, しかし英文の中で最も有名なフレーズが連続してあらわれることになる. したがってまったくでたらめな  $C$ : 参照情報を与えても, 出力には有名なセンテンスが多く出力されることになる.

ここで  $C$ : 参照情報として  $C_1, C_2$  いずれの場合も与えた参照情報次第で任意の文字列を生成し得ることに注意してほしい.  $E_1$  の蓄積情報は均等分布であるから指示  $C_1$  の与え方がでたらめならいかなる単語, 文字列もほぼでたらめに出力する.  $C_2$  の場合は不均等分布であるため, 有名でよく知られたセンテンスがより高い確率で出力される. しかしいかなるセンテンスも確率は低いがそのような文字列を出力する可能性がないわけではない.

DeepDream でも  $C$ : 参照データとして「ホワイトノイズ」を与えることができる. その場合に生成した画像が

Mordvintsev [2]により示されているので図 5 に引用して示す.

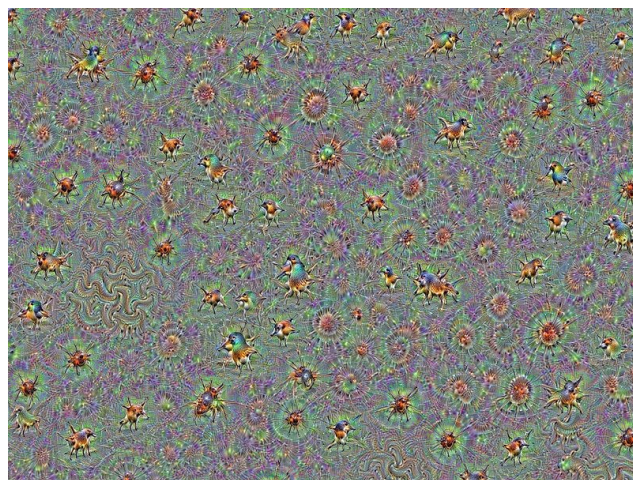


図 5 機械学習による作画例( $E_2$  のみが寄与する場合)  
Mordvintsev [2]による

Figure 5 An example image created by machine learning  
(when only with contribution of  $E_2$  by Mordvintsev [2])

$C$  にランダムノイズを入力した場合  $C$  は単なるノイズであるから  $F$ : MACW の知的財産的価値への寄与はないと考えるべきだろう. このようにして生成した文章や画像の知的財産的価値への寄与は  $E$ : 蓄積情報の寄与によると考えられる.

ここまでの考察をまとめると以下になる.

- (1) MACW には蓄積情報と参照情報が寄与する.
- (2) MACW の評価においてそれらの情報量が関係する.

それでは, 単純な情報量に基づいて知的財産的価値への寄与の評価を行えば適切な評価が可能だろうか. 単純な情報量によって創作物への寄与を評価する上での問題点とその対策について次に議論する.

## 4. AlgIM による評価

### 4.1 アルゴリズム的信息量(AlgIM)

議論の前にアルゴリズム的信息量(以下 AlgIM)の概念を導入する. AlgIM の提案はチューリングにさかのぼることができる. たとえば円周率  $\pi$  の小数点以下の桁の連続は, 一見ランダムなデータで無限の情報量を持つように見えるが, 一定のアルゴリズムにより「計算可能」である点でランダムな数字の列とは異なる. チューリングはこのような情報をあらわす「計算可能数」という概念を提唱した.

G. J. Chaitin 等は, これをさらに発展させ AlgIM という概念を提唱した[8][9][10][11][12].  $\pi$  の桁の並びはほぼランダムであるが,  $\pi$  の生成アルゴリズム自体はコンパクトである. したがって  $\pi$  を記録再生するには無限のメモリは必要ではなく,  $\pi$  のアルゴリズムと計算領域があればよい.



データの情報量を、そのデータを生成し得るアルゴリズムのサイズとして定義すれば、 $\pi$  のような計算可能数の情報量は有限でありアルゴリズムに基づいた情報量を定義できる。

#### 4.2 AlgIM による知的財産的価値への寄与評価の利点

$\pi$  の各桁における数はランダムに出現し、あらゆる数の並びがほぼ均等に出現することが経験的に知られている。したがって  $\pi$  を 27 進法で表現し、各桁にアルファベットと空白を割り当てれば、Simon & Garfunkel, “The Sound Of Silence” の歌詞や、シェイクスピアの小説も、いずれかの桁に出現する。

そこで  $\pi$  を蓄積情報とし、蓄積情報の位置を指定することで出力を行う場合を考える。先の議論において蓄積情報からの情報量を問題にしたが、蓄積情報から直接得られる情報のみが重要であるという立場であれば、この場合出力はすでに蓄積情報に含まれていたのだから、知的財産的価値への寄与はすべて蓄積情報から派生したとみなせる。つまり、MACW への知的財産的価値への寄与は蓄積情報のみということになる。

しかし蓄積情報が  $\pi$  という数値の一部であることを考えると、そのような解釈は不自然とも考えられる。もしこのような場合に蓄積情報が創作の主な貢献要因と考えると、すべての文学作品は  $\pi$  に含まれているので新しい作品ではない、と考えることになってしまう。

このような困難を避けるために AlgIM による情報量の評価を取り入れることに利点があると考えられる。

前述のように蓄積された  $\pi$  の AlgIM はごくわずかである。 $\pi$  のすべての桁が、表面的に多くの情報量を有していても、AlgIM は少ない。したがってアルゴリズムの歌詞の情報のうち  $\pi$  に起因するものは相対的に少なく、大部分の情報は「桁数指定」という指示によって与えられたと解釈できる。 $\pi$  の中に無限のコンテンツが含まれているとしても、AlgIM としては  $\pi$  の計算アルゴリズムの長さに相当する情報量でしかないから、 $\pi$  に帰属する知的財産的価値への寄与は少ないと評価することが可能になる。

$\pi$  と同様に、表面的に多くの情報量を生成する手段は数多く存在する。そのように多量の「表面的」情報を生成することで多くの、知的財産的価値への寄与のある蓄積情報のかわりにとにかく大量の蓄積情報を提供すれば、知的財産的価値への寄与を主張できてしまう。

簡単にいいかえれば、もしも AlgIM を基準としなければすべての創作物は、 $\pi$  の部分列の参照であり、創作性がないという主張が可能になってしまう。AlgIM を基準とすればそのような不合理で不自然な主張を明確に排除できる。

より有益な貢献をした作業者にインセンティブを与える評価とするには AlgIM を基準とすることが有益であると考えられる。

#### 4.3 AlgIM のみによる評価の欠点

それでは知的財産的価値への寄与を AlgIM のみにより定義するのは合理的だろうか。残念ながらそうではないと思われる。AlgIM による評価の欠点として、アルゴリズムが MACW に与える知的財産的価値への寄与をそのアルゴリズムのコードの長さだけで論ずることが適当とは思えない、という問題がある。

AlgIM はアルゴリズムとして最小の情報表現を意味するが、その創作物としての知的財産的価値は AlgIM の大きさより低い可能性がある。わかりやすい例として、まったくランダムな命令列からなる手順を考える。手順自体がランダムなコードからなり、ただし停止したり出力が途絶えたりすることがないように、制約されて作られた、きわめて長大なプログラムであるとする。別のいいかたをすればチンパンジーに作らせたプログラムであるとする。このようなアルゴリズムと同じ出力を得るには全く同じアルゴリズムを実行する他にない。またこのアルゴリズムを実行すればアルゴリズム自身よりも多くの、見かけの情報量を持つ出力が生成される。このようにして「巨大な」AlgIM を有するデータを作ることができる。しかしそのようなアルゴリズムの出力に知的財産的価値があるだろうかと問えば、乱数と同じで無価値だと評価する意見が多数だろう。

AlgIM は乱数発生器の出力のように、見かけの情報量が巨大な場合について、適切な情報量を算定するには有益である。また同種の情報源であれば情報量が 2 倍であれば寄与も 2 倍に評価することには妥当性があると考えられる。しかし異なる素性の情報源の知的財産的価値の比較に AlgIM をそのまま用いるのは適切ではないと考えられる。

#### 5. 情報単価の算定

ここでは、MACW への人による評価が与えられたものとして、それを前章で問題となった異なる素性の情報源の知的財産的価値の比較に反映させる方法の一案を示す。前章までの議論で蓄積情報の評価における困難を回避するために情報量として AlgIM を用いる利点があることを示した。また情報量として AlgIM を用いても、たとえばまったくランダムなアルゴリズムを想定した場合、情報量が巨大であるだけでは必ずしも MACW の知的財産的価値が高いとは限らないことを指摘した。

そこで MACW の知的財産的価値に対する寄与を、情報量という尺度を基本としつつ、情報量毎の寄与率が変化すると仮定したモデルを考えることとする。そのような寄与率の評価をいかにして行えるだろうか。

ここで解決すべき問題点は、MACW の人による評価は可能だが MACW の生成に関与したきわめて多数の寄与の人による評価は困難である、という点である。MACW の知的財産的価値への寄与ひとつひとつは、人間にとっては微細すぎて個々に意味を持たず、それらの寄与を一つ一つ人が

評価することは実際的ではない。

そこで、MACW に対する人による評価から個々の要素の寄与を分解し推定する方法として、線形回帰分析を用いた方法を提案する。

### 5.1 線形回帰分析による情報量あたり情報単価の算出

ここで、参照モデルの C, D, E の構成要素を別個の情報源とし、参照モデルはそれらの別個の情報源から情報を得て MACW を生成するものとする。そして情報源毎に情報量毎の寄与は異なることがあり得るが、同一の情報源の情報量毎の寄与は一定であると想定する。そのような情報量毎の寄与をその情報源の情報単価と呼ぶことにする。

以上の前提のもとで、線形回帰分析による情報単価の算出方法を示す。

$p_1 \cdots p_N$  を  $N$  個の生成された創作物の最終的な評価、 $A$  を  $M$  個の情報源から  $N$  個の創作物を生成する場合にそれぞれ利用する各情報源からの情報量をあらわすマトリックスであるとする。そして  $M$  個の情報源は情報量あたり一定の情報単価、 $r_1 \cdots r_M$  を持つと仮定する。

両者の関係は式 1 のようにあらわせる。ここで  $e_1 \cdots e_M$  は近似誤差である。誤差最小となる  $r_1 \cdots r_M$  を求めることで各情報源の創作への寄与率を推定することができる。

この問題は線形回帰分析となり、データが与えられれば簡単に答えをもとめることができる。その結果各情報源の情報単価を形式的に与えることができる。

$$\begin{pmatrix} p_1 \\ \vdots \\ p_N \end{pmatrix} = A \begin{pmatrix} r_1 \\ \vdots \\ r_M \end{pmatrix} + \begin{pmatrix} e_1 \\ \vdots \\ e_M \end{pmatrix} \quad [\text{式 1}]$$

結果として、モデルに含まれる各情報源について情報単価が算出される。

### 5.2 線形回帰による情報単価の性質

ここでは導出した情報単価がどのような性質を持つかを論ずる。

算定の基礎は MACW の知的財産的価値に対する人による評価である。人による評価はそれ自体がきわめて複雑な問題である。たとえば MACW の利用実績や人による値付けをした場合の価格などが考えられるが、どの方法が妥当であるかについてのコンセンサスはない。しかし、ここでの目的は MACW の人による評価から各情報源の寄与への分解の方法であるから、MACW の知的財産的価値は外部的に与えられるものとし、ここでは論じないこととする。

MACW の人による評価には避けられないゆらぎや誤差が含まれる。しかし提案した方法(線形回帰)では多数の評価を平均することにより情報単価の信頼度を増すことが可能である。すなわち、MACW の入力として利用された様々な情報はくりかえし様々な MACW に利用されることで、その情報単価が「平均化」され、より正確な情報単価が得られる。

情報単価と情報量が MACW の知的財産的価値への寄与の妥当な指標となっているか否かを、式 1 による近似誤差が十分小さいか否かによって検証することができる。

情報単価は情報源毎に評価されるから、競争により動的に変化することも可能である。たとえばある情報源の寄与と置き換え可能という意味で類似のものが増えれば、その情報源は「陳腐化」し、知的財産的価値への寄与が下がる。独自性が高い場合には寄与の評価が高くなる。

以上のように情報単価を考慮した情報量は、知的財産的価値への寄与の指標として望ましい性質を備えている。

## 6. まとめ

本稿では機械学習を利用した「創作物」の知的財産的価値への寄与の評価を情報理論的測度の観点から考察した。

はじめに参照モデルを提示して形式的な議論の準備を行った。第二に、参照モデルに入力される情報量を分析し、情報量が知的財産的価値への寄与に対応する場合と、情報量のみでは知的財産的価値への寄与の評価に問題が生じる場合を示した。第三に、AlgIM を導入し、第二の例の問題の一部が解消されるが、AlgIM でも異なる素性の情報源の知的財産的価値への寄与を適切に評価するには十分でないことを示した。最後に人による評価の導入を提案し、MACW の人による評価から、参照モデルの様々な要素の寄与を分析し、情報単価を推定する方法を論じた。

機械学習による創作は今後爆発的に普及すると予想される。本論文は知的財産権などの法的問題は論じないが、機械学習の素材となる創作物の供給を始め、機械学習を利用した創作に必要なすべての貢献のインセンティブをいかにバランスよく高めるかが、機械学習を利用した創作を発展させるための制度設計において重要であると筆者は考える。

機械学習を利用した創作は多くの構成要素からなり、機械学習のトレーニングに用いた創作物や機械学習に与えるパラメータ、最終的な創作時に参照する参照データなど、多くの素材に依存する。それらの寄与度の評価をどうするかという点は現在議論が続いている点である。

機械学習の利用により創作物の知的財産的価値への寄与の評価に変化がおきることには大きな不安がある。しかし同様の変化は、印刷、録音録画が実現した際に起こった際にもおきたことである。録音が実現する前は個々の演奏が寄与を持っていたが、録音により同じ演奏を再現することの付加価値が下がり、そのかわり複製の元となる高度な品質を持つ演奏者や、個別の条件に合わせたアドリブ演奏ができる演奏者の寄与が上昇した。

音楽配信においても 2017 年の全世界音楽売り上げは 173 億ドルでさらに増加を続けているが、Value Gap、すなわちサービスプロバイダが得る収益と音楽の権利者の収益のギャップが拡大していることが懸念事項となっている[13]。創作方法や創作物の利用方法の変化により創作を衰退させ

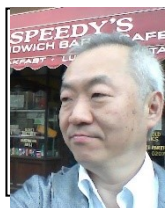
ないためには、そうした変化に遅れることなく適応していく必要がある。

機械学習を利用した創作の活性化には、すべての利害関係者にバランスのとれた適正な収益がもたらされることが重要である。本論考が、機械学習を利用した創作を発展させるための制度設計に活用され、その発展に寄与することを期待する。

## 参考文献

- [1] Alexander Mordvintsev et al. “Inceptionism: Going Deeper into Neural Networks”.  
<https://ai.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html>, 2015/6/17, (参照 2018-06-29)
- [2] WIKIMEDIA COMMONS. “Category: Deep Dream Images”.  
[https://commons.wikimedia.org/wiki/Category:Deep\\_Dream\\_images](https://commons.wikimedia.org/wiki/Category:Deep_Dream_images) edited on 15 May 2018, at 13:20, (参照 2018-06-29)
- [3] Alexander Mordvintsev. “DeepDream - a code example for visualizing Neural Networks”. Google research blog, Posted July 1, 2015. <http://googleresearch.blogspot.co.uk/2015/07/deepdream-code-example-for-visualizing.html>, (参照 2018-06-29)
- [4] Karl, F. Milde, Jr.. Can a computer be an “Author” or an “Inventor”. Journal of the Patent Office Society, 1969, Vol. 51, Page 387
- [5] 金子 格. AI の産業応用の拡大における知的財産の扱いに関する考察. 情報処理学会研究報告, 電子化知的財産・社会基盤, 2015, 2015-EIP-69(8),1-4
- [6] 金子格. 人工知能による著作物の創作性尺度に関するアルゴリズム情報理論から見た考察. 2017, 2017-EIP-76
- [7] 金子格. 機械学習技術による創作物の知的財産評価のための情報理論測度について. 情報科学技術フォーラム 2018, FIT2018 CN-001
- [8] Toby Haddad, Gregory J. Chaitin. Randomness and Mathematical Proof, Scientific American. 1975, May, pages 47-51
- [9] Gregory J. Chaitin. Algorithmic Information Theory. IBM Journal of Research and Development, Vol. 21, Issue 4, 350-359 (1977)
- [10] Ian Stewart. The ultimate in undecidability. 2015, Nature, vol. 332, pages 115–116
- [11] Gregory J. Chaitin. Randomness in Arithmetic and The Decline and Fall of Reductionism in Pure Mathematics.1993, arXiv, arXiv:chao-dyn/9304002, RC-18532
- [12] Juergen Schmidhuber. On Learning to Think: Algorithmic Information Theory for Novel Combinations of Reinforcement Learning Controllers and Recurrent Neural World Models. 2013, arXiv, arXiv:1404.7828
- [13] 安藤和弘. 音楽のインターネット送信と Value Gap 問題, 論研ジュリスト. 2018, No 26, p.20-27

## 著者紹介



金子格（正会員）

名古屋市立大学客員准教授, f2f computing 技術顧問 早稲田大学 (1980), 博士(情報科学 2002), 日立、アスキー、GCL、東京工芸大学を経て 2018

より現職

<< 印刷対象外 >>