

小説の単語分散表現に基づく機械学習評価手法の提案
--オンライン小説投稿サイト「小説家になろう！」作品群を中心に--

国際日本研究講座
C2KM1006 王劭宇¹

1. はじめに

本研究は機械学習モデルによって、人気小説の共通パターンを検出することを目標とした。人気小説は、叙述技法、題材、登場人物などの側面において、何らかの共通点が存在することは経験論によって導き出せる。それは数多の物語創作ガイドブックから窺える。その中に、最初はハリウッドの脚本家で流布し、後世界中を広まり、ビジネス小説の基準になったのがキャンベルの理論である。

キャンベル（2015）の「英雄の旅」やスナイダー（2010）の構成法は抽象性が高く、解析することに活用することには制限が多くなかなか困難である。まだしも複雑のゆえに読者の認知反応と直接に繋がれず、交互作用を説明するのがもどかしく考えられる。一方、物語論においてすでにジュネット（1985）が時制と態などの解析基準を唱えた。しかし、ジュネットの理論によって物語の諸技法が分類されたが、人気作品のパターンまで言及されなかった。

但し、これまでの試しは経験則の範疇から外さず、個人の経験と解釈に左右されがちである。つまり、世の中にあるすべての作品を分析して得られた結論でもないし、都合よく解釈で無理やり当てはまらせることもできる。主観性に依存するのである。一方、人気とは集団的心理に依存するものであり、個人は膨大な経験を積まない限り、個人の感覚と知識で推測することが難しい。仮に膨大な経験によって共通されたパターンがあっても、その存在を証明することが難しい。但し、物語についての理論はほぼ曖昧であり、どの作品に対しても簡単に当てはまる。理論と実務の間に深いギャップがある。果たして人気作品において共通したパターンがあるか。条件次第でブレが多いため、本研究によって明らかにさせたい。

本研究はオートエンコーダーという機械学習モデルによって、人気作品の特徴を整理し、未知の作品の自動選別を試みた。人間と違い、人工知能には説明性を有しない。つまり、人間が実際の区別をモデルに直接伝えるではなく、情報と結果をモデルに与え、自ら帰納法によって因果律を導き出すのを期待するだけなので、何を以て推論が成り立

¹ メール：shimoha0527@gmail.com

つか、作り手でも知る術がない。ブラックボックス問題²と言われるのである。ただ、途中が明らかでなくても、結果だけ確認できれば、方法の限界と信憑性がはっきり分かる。

ここまで言っているパターンというものは、題材、登場人物、叙述などの側面を渡り、単一のものに定義されかねるが、人工知能の特性を踏まえて、あえてパターンの本質に触れず、未知数 X のような概念的な存在として捉え、存在するか否か、存在する場合、どんな形態で存在するかを論じてみる。

現象の本質に触れず、現象の存在だけに注目するのを、突然に聞いたら不思議と思うかも知れない。それは、後述した仮説を解明するためである。そこで、問題のレベルが一個上（抽象的）違ってくるのである。

なお、純文学作品はそんなに簡単に括ることができないため、本研究はすべてビジネス娯楽小説の話に限定される。

2. 仮説

1. ある条件を満たさない作品は、人気小説になる可能性がない。条件を満たす作品は、人気小説になる可能性を有するが、必ず成り上がることが決まらない。（十分必要条件³ではない。）

2. 条件を満たすか否か二択のみあるため、条件の合致する具合で作品を評価することができない。（分類問題であり回帰問題ではない。）

ここでは「ある条件（パターン）」の中身を探求しない。人気小説と他の作品の中から、分離可能の割合に着目し、条件の条件、いわゆるメタ条件（メタパターン）だけを明らかにしたい。そうすることで、従来の経験論が裏付きされると思う。

仮説はエントロピーの概念で説明される。人気が集うこと自身が高度複雑な現象で、一つの条件でも揃わなければ失敗してしまうため、人気作品は一致性が高い一方、他の作品はバリエーションが無限に多いと推測される。言い換えれば、人気ではない作品に何かがあるから人気がないではなく、何かがないから人気にならないという思考である。本研究はその現象に注目し、一致性によって固定パターンの存在を確認する。そのため、人気作品の具体的な特徴に触れずにいけた。

3. 研究対象

² モデルはデータセットを帰納し、既知の事実から未知のデータを推論する。しかし、どんな思考を沿って結論に至ったか分かる方法がない。複数の高次元の行列をモデルから抽出することが可能にも関わらず、それを意味に持たせる作業自体は人間にとって不可能なのである。モデルの内部は見られない故にブラックボックスと例えられる。結局、入力と出力の対応関係でモデル内部の状況を推測するしかできない。

³ A が成立すれば B が成立する。但し、逆は成立しない。

本研究は小説投稿サイト「小説家になろう！」⁴に掲載された作品群を中心とするオンライン小説が対象となっている。「小説家になろう！」は株式会社ヒナプロジェクトによって2004年から運営されてきたオンラインサービスであり、そこで無料の小説掲載と小説閲覧ができ、オンライン小説サービスの中で一番のシェアを占めている。現時点で⁵は計939403本の作品が掲載されている。「小説家になろう！」以外に、オンライン小説において角川集団が運営する「カクヨム」などのサービスも存在する。

2004年当時から個人非営利で運営されているため、2015年株式会社ヒナプロジェクトが成立したにも関わらず、一向に作者を干渉しない方針を取っている。「小説家になろう」発の小説は、書籍化されたり、更にアニメ化されたりするシリーズが多いものの、作者が出版社の要請によってもともと掲載された作品を削除したり、また内容と題名を大幅に変更したりするため、サイトで掲載されたバージョンと書籍化されたバージョンを同一視してはいけな⁶い。

「小説家になろう」は登録、投稿、閲覧が無料でできることが特徴である。投稿を利用者に任せ、作品や作者へのブックマーク、評価、コメントができる。運営方法から見れば、ツイッターやインスタグラムを含めた利用者生産式メディアサービス

(CGM)に分類される。サイト内のあらゆるサービスは基本的に無料であり、広告収益で運営される。その中、人気の掲載作品が出版社の青眼を受け、作者がデビューすることが毎年の恒例になっている⁷。

オンライン小説の領域の筆頭にも留まらず、「小説家になろう」は紙本書籍、アニメ、映画、コミックなどの媒体を通して大衆に広く知られている。例を挙げれば、佐藤勤「魔法科高校の劣等生」(KADOKAWA、2011年)、長月達平「Re：ゼロからの異世界生活」(KADOKAWA、2014年)、住野夜「君の膵臓を食いたい」(双葉社、2015年)など、いずれもアニメ化⁸され、知名度が高い作品が該当作品の中に存在する。

存在感があまりに強く、「なろう系」という言葉さえ作られている。言葉自体は「平凡な主人公が突如放り込まれた異世界で活躍する」といった安易なストーリー展開が特徴⁹として揶揄で使われるが、「小説家になろう」がすでに大衆の視野に入っていることを示している。「魔法科高校の劣等生」など「小説家になろう」発の作品が広く受け

⁴ <https://syosetu.com>

⁵ 2023年2月22日

⁶ 『『SAO』はもともオンライン小説で無料で読んでいました。しかし、今（書籍化してから）はお金を払わなければ読めません。それを良しとさせる+αの価値をいかに出せるか』（三木一馬「面白ければ何でもあり」株式会社KADOKAWA、2015年、p.168）

⁷ 出版社の要請を応じて運営側が代わりに作者を連絡するだけで、料金を徴収しない方針である。

⁸ 「君の膵臓を食いたい」は映画化も受けた。

⁹ 「デジタル大辞泉プラス」小学館

入れられていることも事実である。

なお、本研究は合計ポイント数の高い 2000 作を人気作品にし、他の作品と区別する。それは単に訓練に適当なデータセットサイズを選ぶことで、境目を移動することで結果が変化するかも知れないが、未来の研究に託しておく。少なくとも 95 万作の中に、2000 作は僅かな 0.2% を占めて、どれも精鋭中の精鋭と言える。そのため、人気作として 2000 作の代表性を認める。

4. 研究方法

4.1. 単語分散表現の取得

4.1.1. 単語分散表現とは

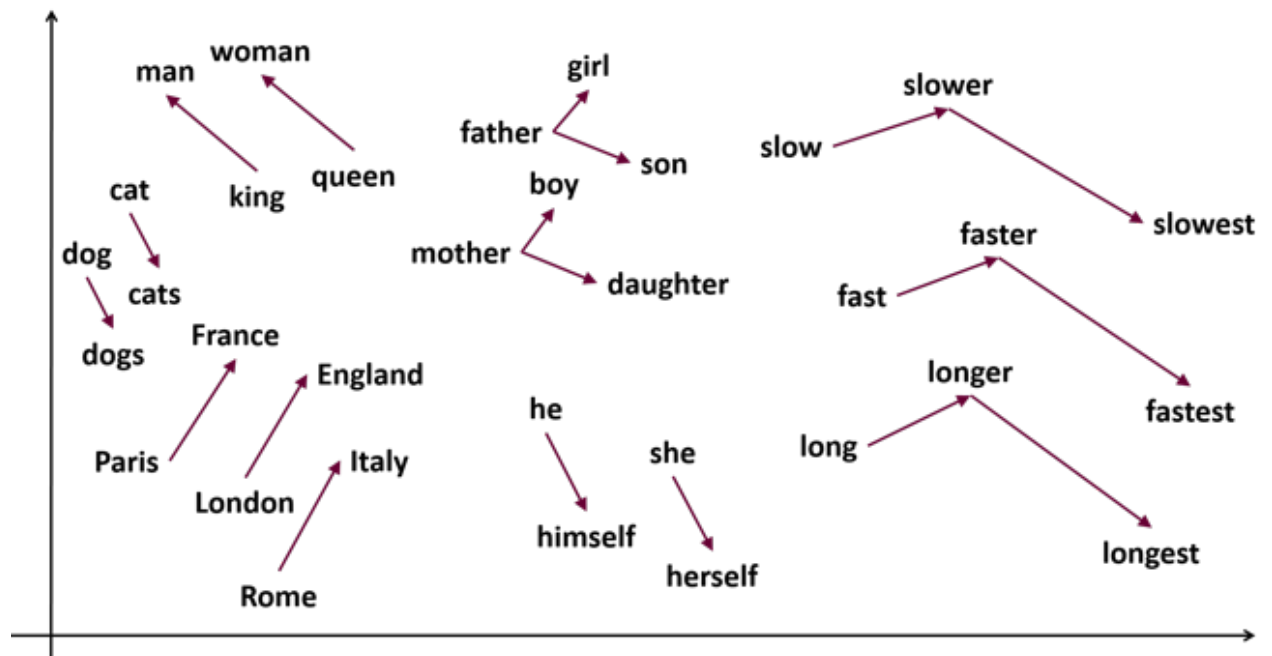


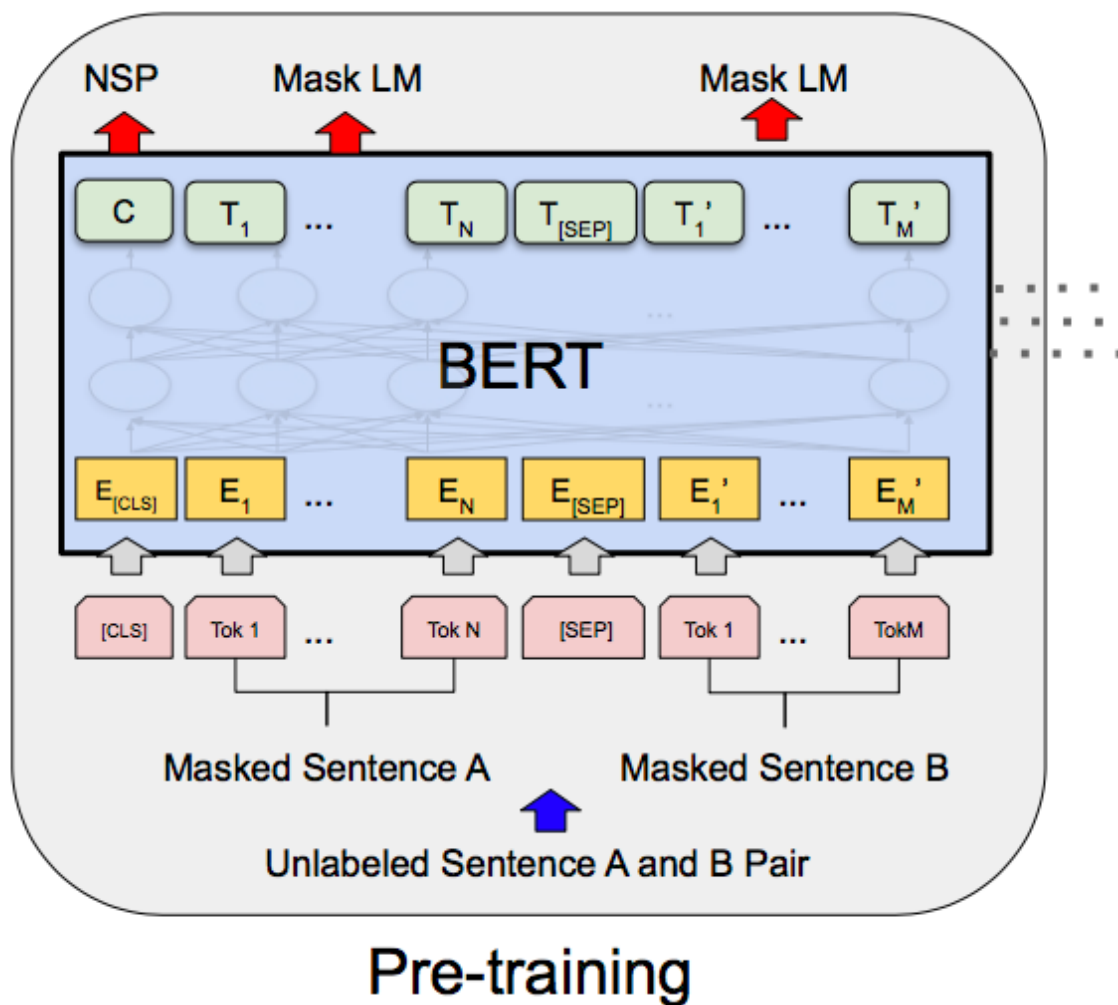
図 1、仮想空間での座標が変われば、意味も変わる。¹⁰

単語分散表現とは、単語が内包する意味の表し方である。語彙は記号としてシニフィアンとシニフィエの二つの側面があり、文字や音声のみを取り上げれば、ただの記号になるため、意味との連結があるからこそ、コミュニケーションが取れる。しかし、意味は話し手と聞き手の脳内にあり、観測や表すことが難しいである。辞書のように、単語によって単語を定義することが普通であるが、主観性と曖昧性が強く、意味の表し方として不自由なところが多いと考えられる。一方、単語分散表現は大きいデータベースが裏付けで、客観性と再現性が保証され、性質上曖昧性が少ない。ちなみに、辞書はどうしても脳内思考を記号レベルまで下げないといけませんが、単語分散表現は主に機械学習モデルの中身を抽出する方法にされるため、モデルの思考過程を覗くことができると

¹⁰ <https://samyzaf.com/ML/nlp/nlp.html> アクセス：2023.10.01

も言える。

単語の共起性によって遠近関係を計算し、それを再現されるよう意味ネットワークを構築できる。意味ネットワークにある相対的位置の通りに仮想空間で単語が配置されたものが単語分散表現と言われる。遠近関係は方向性がないため、単語分散表現に方向性があるものの、あくまで相対的な位置で推測されたものであり、軸に特に意味がない。しかし、軸の方向に集まる単語の共通点で、その方向の特徴が窺えられる。ちなみに、仮想空間は次元が高ければ高いほど意味ネットワークにおける再現性が優れる。例えば、ABC がそれぞれ他の単語より 1 単位ほど離れた場合、1 次元の線分上では 3 点同距離の実現が不可能であるが、2 次元の平面では実現できる。その特徴で、単語における意味を仮想空間にある位置で表すことが可能にあり、意味の変化が位置の変化に投影することができる。座標をベクトル¹¹によって表示すれば、複数の特徴が比率的に混合されたものとして定義され、四則計算などのいろんな操作ができる。そうしたことで、自然言語処理のタスクでも、場合次第で画像認識の技術によって解決されるのである。



¹¹ 複数データ（特徴値）の集まりである。ここでは、入力の一つの単語に対する出力が一つのベクトルに該当する。

4.1.2. BERT

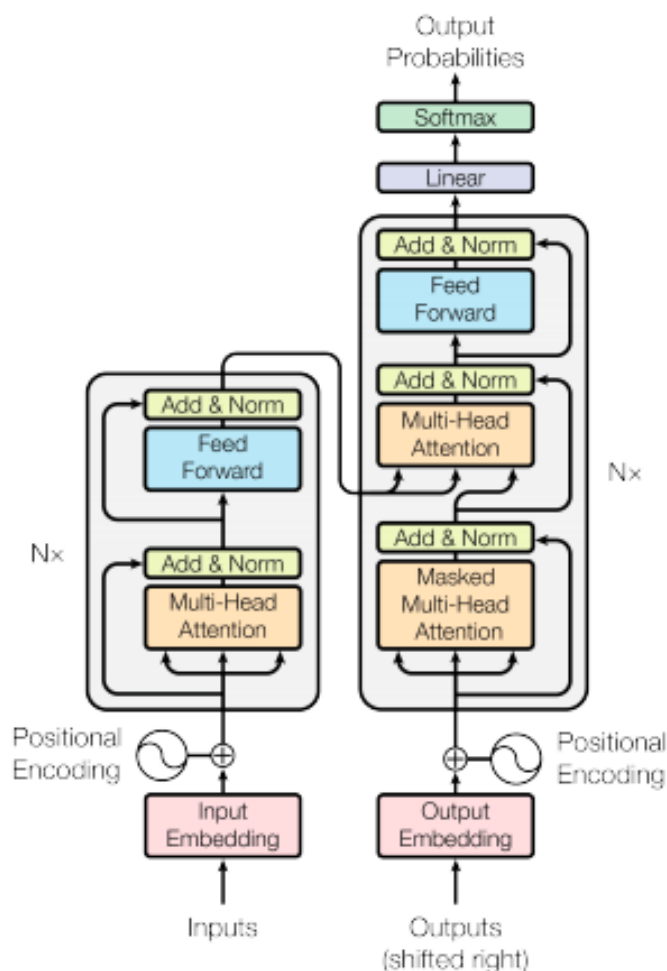


図 2、BERT の構造 (Delvin, 2018) 図 3、transformer の構造 (エンコーダのみ)

本研究は BERT の事前訓練モデルを利用して対象作品より単語分散表現を抽出する。BERT (Delvin, 2018) は事前にオンラインにある膨大な情報とウィキペディアにより訓練されている巨大自然言語モデルであり、幅広い単語の意味を網羅すると思われる。図 2 の示した通り、BERT は部分のマスクされた文に対して復元するタスクと、文の前半の入力によって後半を推測するタスクが課せられ、訓練されている。従来の Word2Vec モデルは、目標の周囲に発生する単語によって単語分散表現を定義するという分布仮説に基づいたものである点が BERT と同じであるが、BERT は transformer の構造¹²が内包されるため、文脈にある他の単語の位置でも考慮に入っており、文章を総合的に理解することができる。なお、巨大な訓練データのおかげで、BERT は場合によ

¹² 正確に言えば transformer のエンコーダ部分のみ内包するが、混乱を招かないように省略することにした。

って人間並みあるいは以上のパフォーマンスを達成することもある。

BERT から単語分散表現を取得する方法は多様であるが、本研究は BERT の最後の隠れ層アウトプットを単語分散表現とするという一番普遍な方法を採用する。そこで、最終ブロックにある transformer 構造から文における各単語の貢献度が抽出される。それを単語ごとの分散表現にかけることで、文ごとの単語分散表現が算出される。

BERT は図 3 の示した通りの transformer を複数重ねる形になっている。単語はリストから番号を振られ、モデルに流れ込む。入力 transformer を通りつつ、ベクトルとして変わっていき、最後入力と同じく形に還元される。層と層の間で渡されるベクトルはモデルの言語とも思われ、ベクトルの変化がモデルの思考過程を表す。すると、最終隠れ層の出力が一番目標らしい数値に整形され、次の出力層に渡され、リストの番号に戻り、さらにリストで自然言語に転換される。そのため、最終隠れ層の出力を単語分散表現に採用することにした。

4.2. オートエンコーダーによる予測

34-layer residual

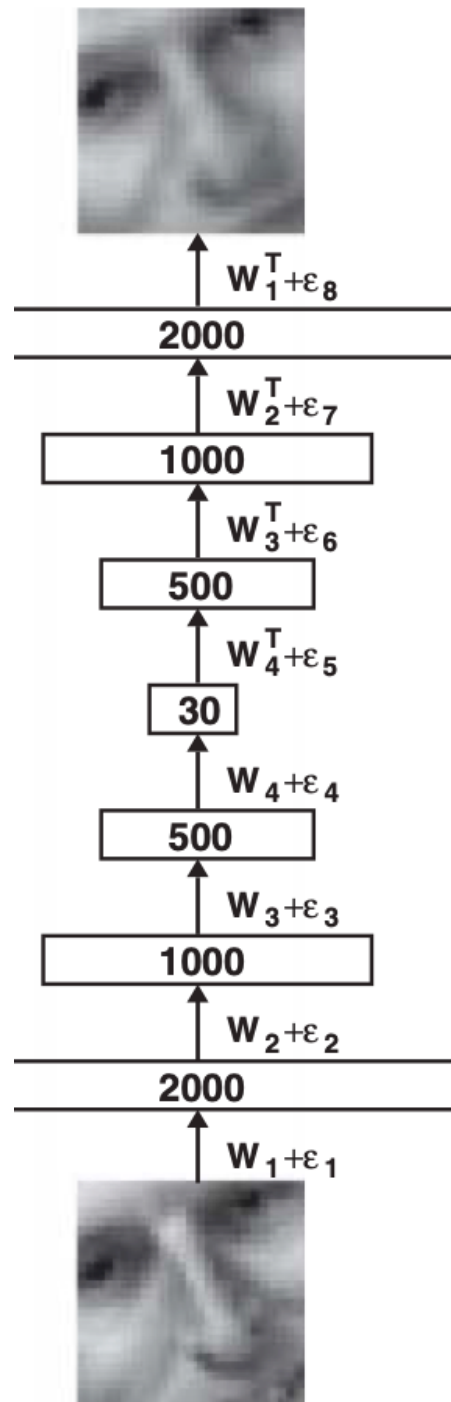
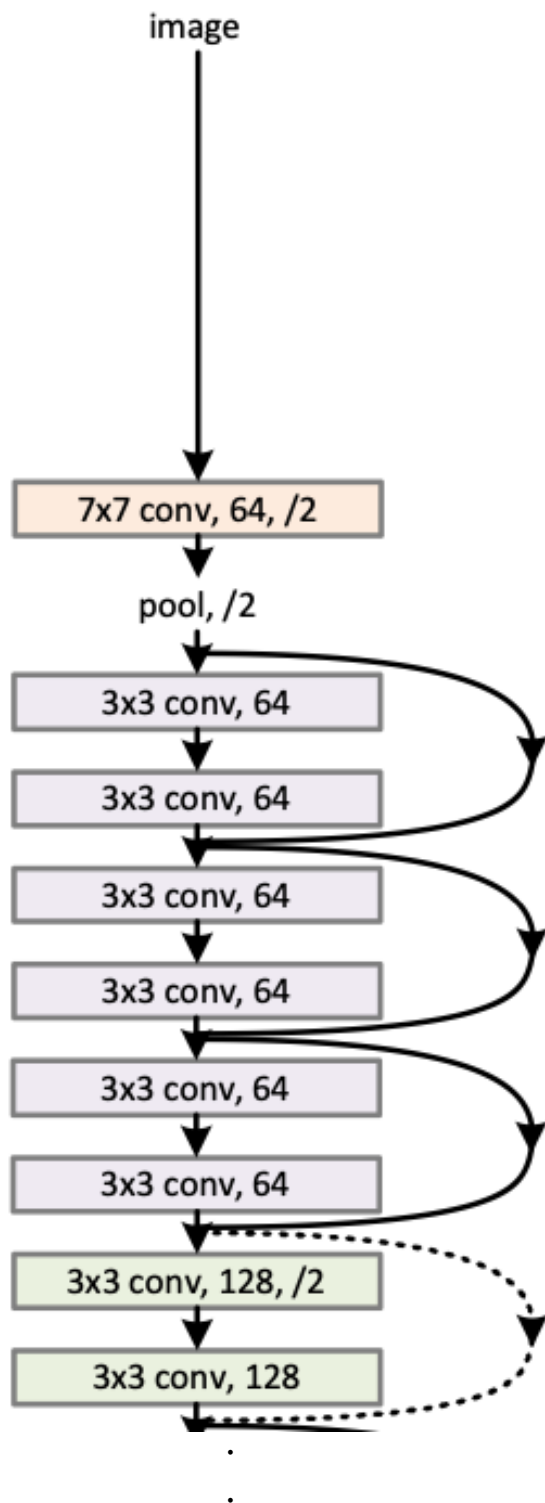
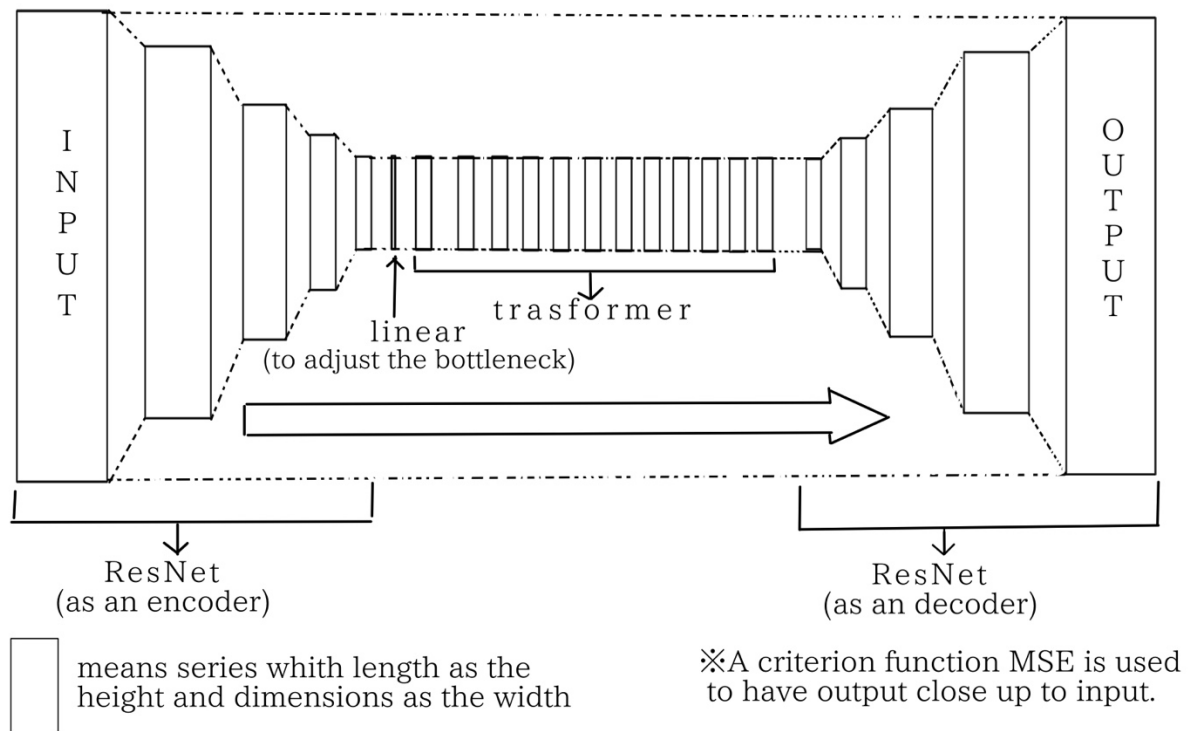


図 4、ResNet の構造 (He, 2015) 図 5、オートエンコーダの構造 (Hinton, 2006)

Auto Encoder Used for the Experiment



ここまで本文データが文字列であり、計算できないため、文単位で区切りし、BERT Tokenizer によって単語分散表現を文ごとに取得する。単語分散表現によって物語の展開と伴いに発生した意味の変化が数値化され、画像認識の手段で分析可能にされる。

オートエンコーダは元々異常検出が一番普遍的な使い道であるが、本研究はそれを逆転し、正常検出に使う。異常検出の場合、オートエンコーダは大量の正常データで訓練することで正常データを認識させるようにされる。すると、異常データに対して認識不可能になり検出できる。しかし、それは「正常データ」が有限パターンでしか通用する。例えば、網目に穴が空いたり、金属パイプに錆があったりすることが検出できる理由は、網目と金属パイプの写真はパターンが少ない。それさえ認識すれば、錆や穴はどんな形になっても、とにかく通常パターンと違うため認識される。言い換えれば、オートエンコーダはエントロピーの低いものだけ適用できる。

一方、小説作品はそうではない。ほとんどの小説は人気がない。むしろ世に出る小説は大抵異類と言っても過言ではない。すると、すべての作品からパターンを帰結しよう

としても、パターンが無限にある結論しか辿り着かないはずである。人気の小説に共通点が少ないのである。それはオートエンコーダの低エントロピー原理を反する。そこで、割合から見れば異類の人気作品群を「正常」、その他の作品を「異常」と定義し、逆の構図にする手入れをせざるを得なかった。ここでは「正常検出」と命名する。母集団の分布による仕方のない制限なので、統計上の正解率の低下に繋がることもある。それについて後述するため、ここで省いておく。

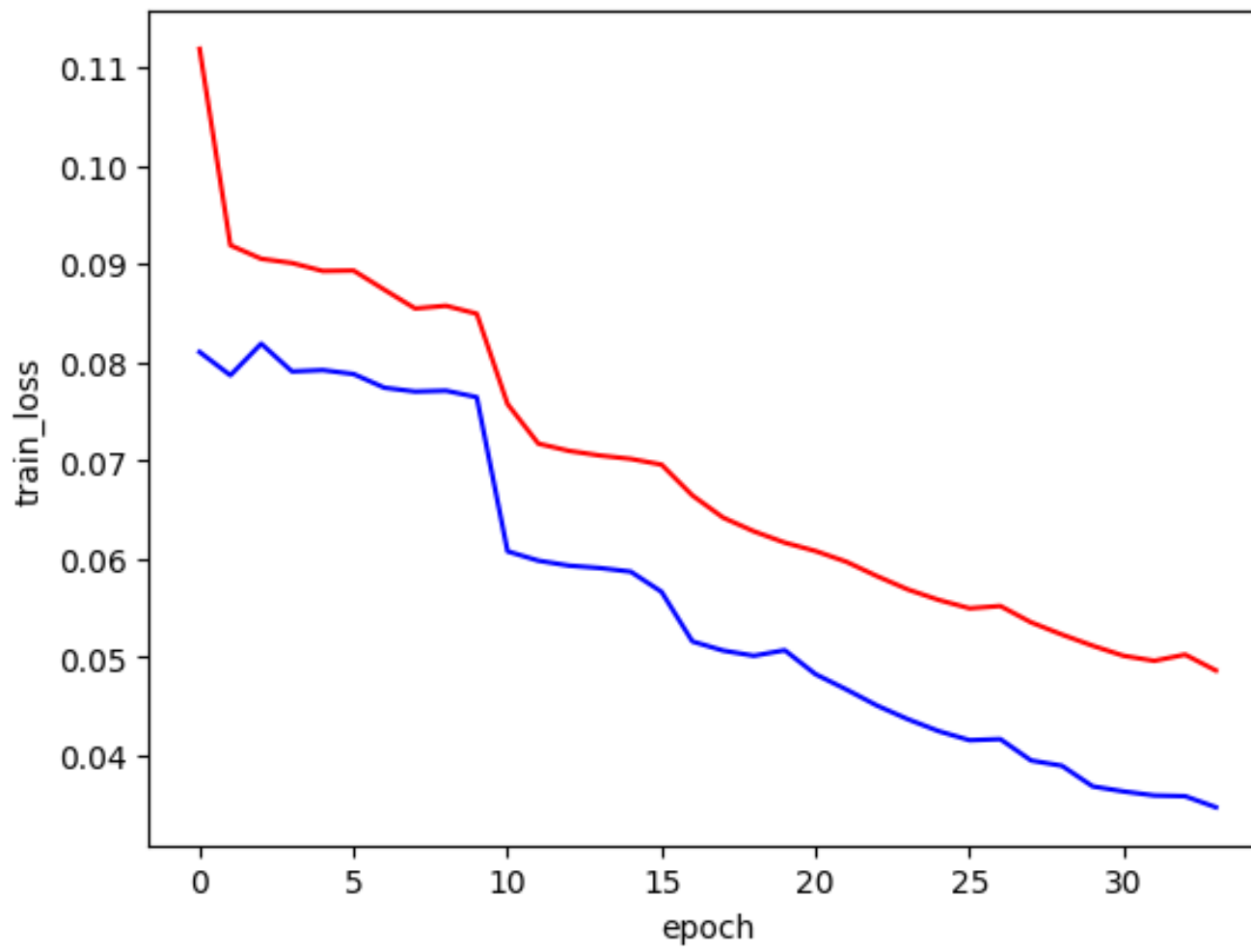
オートエンコーダ (auto encoder, 自己符号化器) はヒントン (2006) が提案した構造であり、通信境界の同名装置をモチーフとした。オートエンコーダはエンコーダ (encoder, 符号化器) とデコーダ (decoder, 復号器) の二つのパーツに分かれ、ボトルネックを作ることによって情報のパターンを収束させる装置である。情報の表現可能性はデータの次元数と比率になる。情報はエンコーダを通せば次元数が削減され、デコーダを通せば次元数が復号される。そうすれば、ボトルネック部分で情報は不要な特徴が削除され、必要な特徴のみ残留する仕組みになっている。何の特徴を残して、何の特徴を消すかはモデルがデータセットによって学習する。すると、共通しない特徴は出現率が低いため、モデルに削除される傾向がある。挙句、訓練データの共通するパターンだけ復号でき、その他のパターンがあっても認識されず復号できなくなる。正常データに対して復号する際の損失が低く、異常データにのみ損失が高いわけである。しかし、正常データにも認識できるパターンがない可能性と異常データにも認識できるパターンがあるため、第一種過誤と第二種過誤 (後述) の発生が問題になる。それはボトルネックの狭さ (認識されるパターン種類) を調節することで調整した。

エンコーダとデコーダはただ鏡像する場合が多いが、内部がどんな構造にされるか決まらない。画像認識の場合、畳み込みニューラルネットワーク (Convolutional Neural Network, CNN) を組み込む場合が多いである。但し、CNN は巨大化する場合、勾配消失¹³が発生することが多いため、さらに ResNet¹⁴ (He, 2015) の構造を採用するにした。なお、作品は転換されたにも関わらず、長さが決まらない可変長データになるため、可変長データを取り扱うことができる transformer をボトルネックの部分に仕込むことにした。そうしたことで、長さに関係なく解析することが可能になる。ResNet と transformer の採用で、高い精度が期待される。

4.2.1. 学習経過

¹³ デープランニングは複数の層を深く重ねた方法であるが、最初の情報は層に計算して次に渡されるたびに段々歪んでいき、最後情報が消えて深い層が学習できなくなる現象。

¹⁴



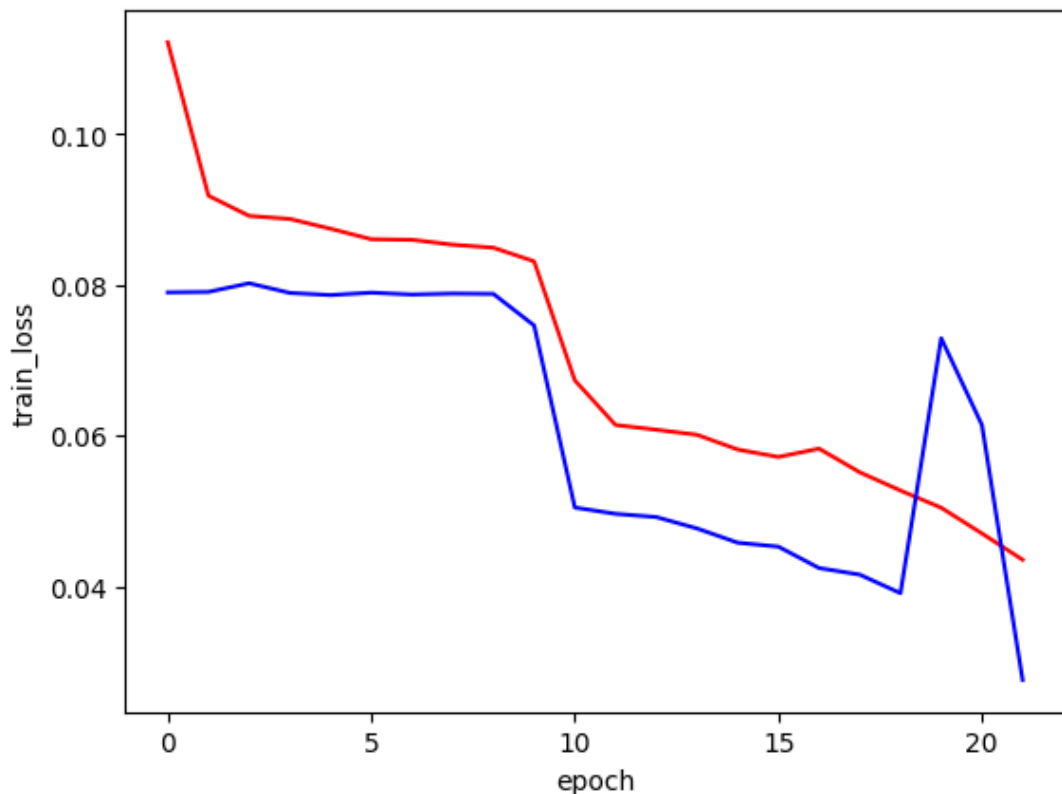


図6 損失の効果経過（最低次元数：12（左）、48（右））

図6によって、損失が降下し、学習が進むことが分かる。但し、どっちもそこに打っていないため、学習回数を増やすれば、さらに精度が出る可能性がある。ここで損失とはモデルのアウトプットとインプットの差である。つまり、損失が低ければ低いほど、モデルの予測が現実には当てはまると思われる。赤は学習損失で、青は評価損失である。機械学習モデルに過学習という答えを当てはめる癖があり、訓練データだけに高精度で動き出すが、実際のデータに参考できるほどの結論に至っていない。それは訓練データに偶然に揃った関係ない特徴も取り出されたためである。それを塞ぐべく、Valid Dataset という部分のデータセットを切り出して、学習させないようにし、過学習が起きたか否かを監視した。図面から見れば、青線（Valid Dataset に対する評価損失）が赤線（学習損失）、学習が正しく進んで、過学習が起きなかったことが分かる。最後の損失は両方大差がないので、次元数が低い方（左）が効率よく分類できると考えられる。

5. 評価

5.1. 基本統計量

評価は訓練データセットとそれ以外のデータからそれぞれ同じ数のサンプルを無作為で抽出した。以下、便宜上訓練データセットから抽出された人気小説のサンプルを正

常データといい、それ以外の人気のない小説のサンプルを異常データということにした。(あくまで便宜上のため、人気のない小説が異常に意味しない。ただ異常検知の流れを合わせる仮の名前である。) 正常データと異常データをモデルに流れ込み、モデルの復号に当たった損失を記録した情報が評価することに利用される。

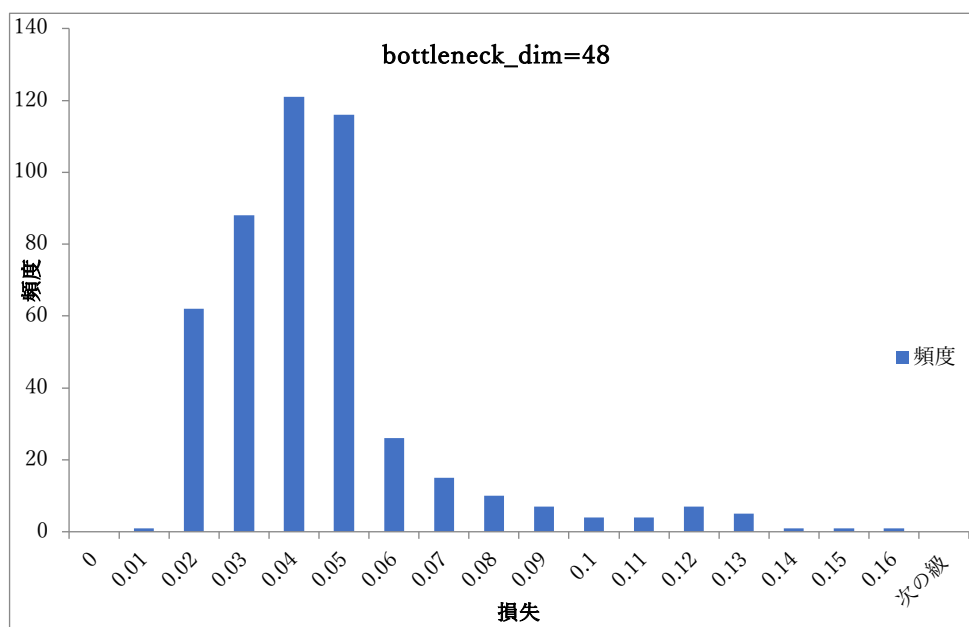
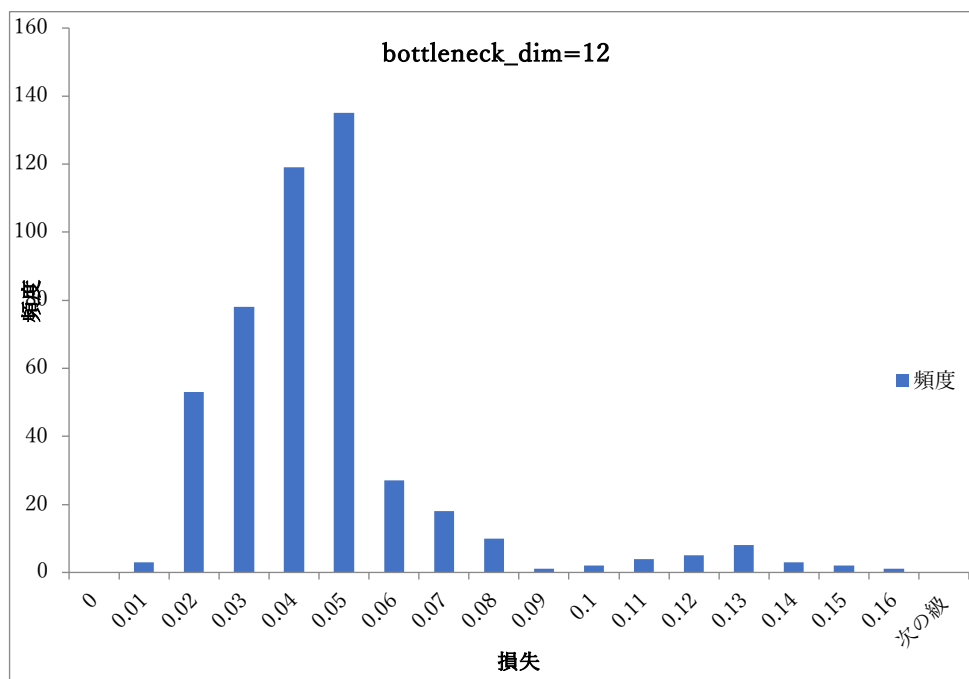


図 7、正常データに対する損失分布ヒストグラム (モデル最低次元数:12(左)、48(右))

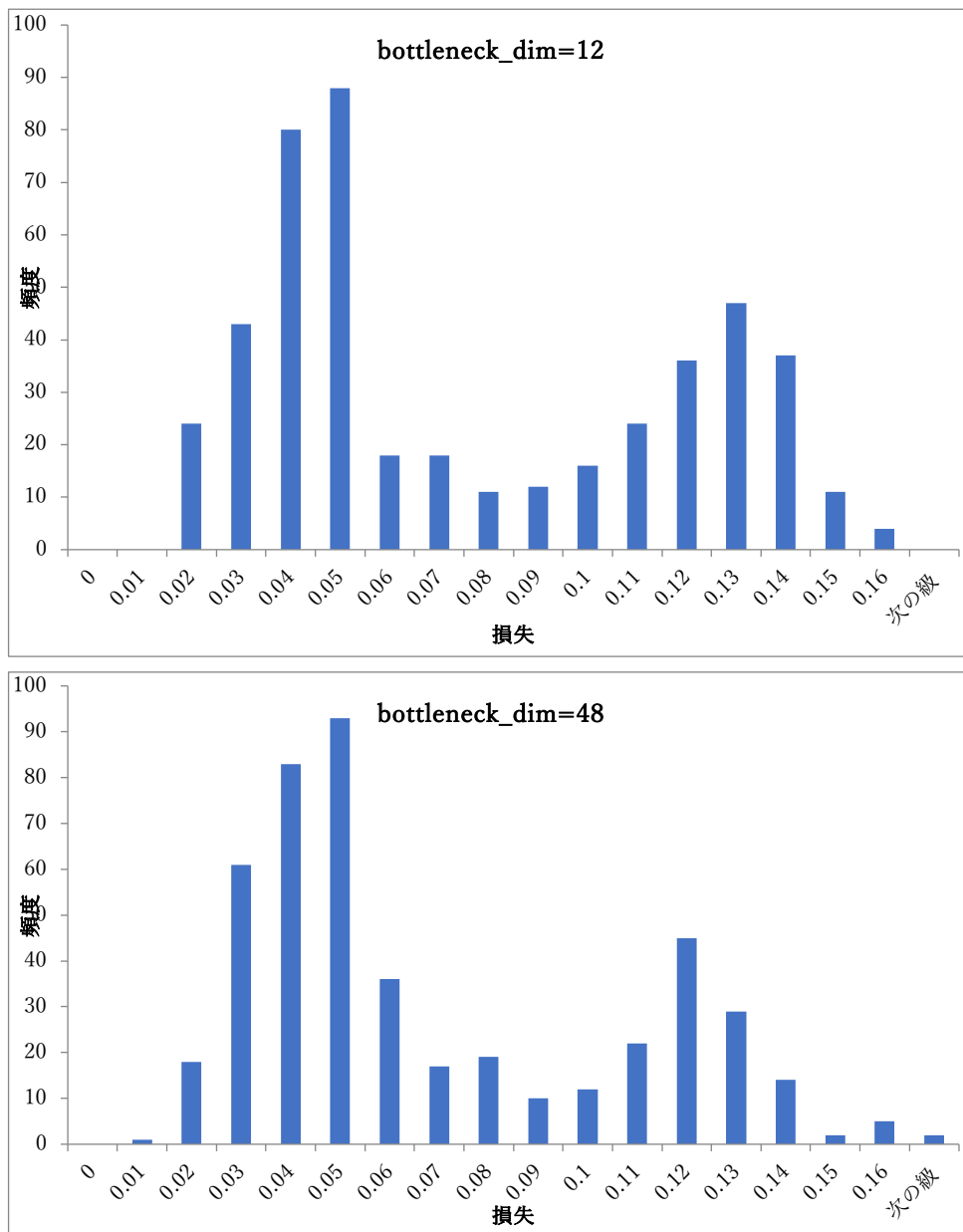


図 8、異常データに対する損失分布ヒストグラム (モデル最低次元数:12(左)、48(右))

図 7、図 8 から正常データと異常データに対する挙動が違うことが分かる。正常データはピークが一つだけあり、異常データは右にピークがもう一つがある。オートエンコーダは正常データに対して低い損失で符号できるが、訓練データセットに入っていない異常データに対して、同じ精度で復号されるサンプルもあるが、復号されず損失の高いサンプルも多くあることが右にあるピークから読み取られる。正常データに対する判定は信頼度が高い一方、異常データに対する判定が不完成である。モデルが上手く学習できていれば、正常データが左側にピークとなり、異常データが右側にピークとなり、ピークが離れば離れるほど、判定が上手く効くはずである。図 7 において、どのモデルも同じ挙動をしたが、図 8 において最低次元数 12 モデルのほうがより高い損失が起

きる傾向にあるため、最低次元数 48 モデルより差が明らかで判定が効くと思われる。

そこで、閾値をどこに設定するかが問題になる。実用することにあたり、閾値より低い損失が起きるデータを正常データに判定し、より高い損失が起きるデータを異常データに判定しなければいけない。本研究のモデルの場合、正常データと異常データのピークが 0.5 より左にあるため、0.5 以下ができれば人気作品群の共通パターンに反応したと推測できる。0.5 以上ができるのが異常データのみの特徴と見なす。なお、異常データはヒストグラムから見た左ピークにある場合、そうしたことで正常データに誤認されるが、本研究の方法の限界で解決できないと推察する。正常データにも僅かな誤認される可能性が存在するが、大まかに正しく認識される。

5.2. T 検定

	unnormal	normal
平均	0.07117982	0.04142757
分散	0.00176251	0.00055432
観測数	469	469
仮説平均との差異	0	
自由度	736	
t	13.386273	
P(T<=t) 片側	4.9377E-37	
t 境界値 片側	1.64692659	

表 1、最低次元数 12 モデルの T 検定結果

	unnormal	normal
平均	0.06350425	0.04060049
分散	0.00143163	0.00051047
観測数	469	469
仮説平均との差異	0	
自由度	764	
t	11.2553112	
P(T<=t) 片側	1.3396E-27	
t 境界値 片側	1.64685052	

表 2、最低次元数 48 モデルの T 検定結果

T 検定とは二群の差をランダム抽出による誤差を仮定し、その可能性が十分小さいかどうかを計算する検定である。表と表はどちらも t 値が境界値より小さく、両群の差

は 95%の可能性で実際に存在することを表し、さらに P 値から間違い可能性が 0.01 以下にあり、顕著性がかかなり高いことを表した。T 検定によれば、どのモデルも正常検出能力が確実にあることが示された。なお、T 検定は違いを検測する検定であり、正常データに対する損失と異常データに対する損失が違うことに裏書きしたが、検定統計量でモデル性能の優劣を判断できない。

5.3. パフォーマンス

判定\事実	最低次元数 12 モデル				最低次元数 48 モデル			
	normal		unnormal		normal		unnormal	
normal	388	83%	235	50%	388	83%	256	55%
unnormal	81	17%	234	50%	81	17%	211	45%
合計	469	100%	469	100%	469	100%	469	100%

表 3、判定結果

表をよって、異なるデータに対して、モデルの判定結果を表した。それで正解率を直接計算できるが、データの次第でそれぞれ性質が異なるため、正解率だけ注目すると、問題の所在を見失いやすいのである。

最低次元数 12 モデルの場合、真陽性：83%、真陰性：50%、偽陽性：50%、偽陰性：17%と表より示される。正常データは異常データより正しく判定されることが多いと見られる。偽陰性が低いため、第二種過誤（Type II error, α error, ほんやりものの誤り）が発生しにくいことが示された。しかし、偽陽性が異常に高いため、第一種過誤（Type I error, β error, あわてものの誤り）が頻繁に発生することが示された。言い換えれば、該当モデルは正常と違うデータでも正常に判定してしまう傾向があり、異常データに弱いである。真陽性と真陰性を合わせ、正解率が 66%だと分かる。それにも関わらず、実際の正解率は真陰性 50%まで落ちると推測する。というのも、正常データは全体の 1%未満のためである。ほとんど異常データしか判定しないため、異常データに弱い特徴が拡大するのである。

最低次元数 48 モデルの場合、真陽性：83%、真陰性：45%、偽陽性：55%、偽陰性：17%と表より示される。正常データに対する挙動は同じであるが、異常データに対してさらに弱いのである。真陽性と真陰性を合わせ、正解率が 64%だと分かる。それにも関わらず、実際の正解率は真陰性 45%まで落ちると推測する。

以上の計算を通して、判定されたデータによってモデルのパフォーマンスが違ってくることが分かった。なお、オートエンコーダモデルの最低次元数を低くすることで、異常データに対するパフォーマンスを上げることが可能であることが確認された。

5.4. 人気と損失の相関性

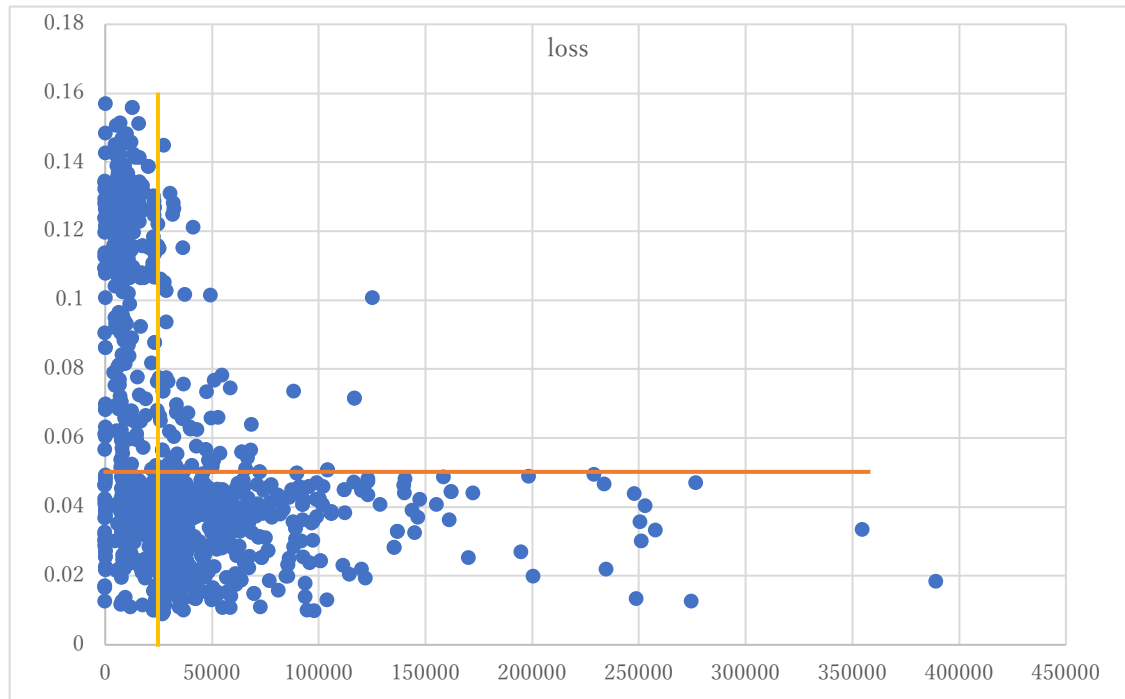


図 9、合計ポイント数-損失 yyplot

ここまでの分析によって、異常データごとと人気のない作品に対する判定が緩く、異常データでも通させるあわてものの誤りが発生することが問題視される。精度を上げるためにはオートエンコーダのボトルネックを狭くし、最低次元数をもっと小さく設定が有効であるが、改善幅が小さく、調整する余地が少ない。最低次元数が小さすぎると、正常データすら判断できなくなり、存在意味も失うのである。もう一つの道は正常データの定義を見直し、正常データに判定する閾値を調整することである。但し、再定義しても改善する余地が小さいと推測される。限界の原因は、仮説の一条目と合致する。

仮説によれば、人気作品に共通するパターンが存在するが、逆は決まらない。つまり、そのパターンは人気のない作品にも散見する。むしろ共通パターンが果たして存在するとしたら、小説の基本となるものになるため、普遍的に作品に存在するはずである。それにしても、他に何か足りないか、単に運がついてないか、人気にならない理由が分からないが、人気になる条件を持っているものの、どうしても人気にならない作品が数多にある。結局、人気作品と認められる閾値を下げて、それら人気になる条件を持っても人気にならない小説をナワバリに入れないのである。

図 9 は横軸がポイント数であり、縦軸がモデルの復号で発生する損失を表す。人気作品を定義するのは、縦線（黄色い線）を引き、左を異常データに、右を正常データにあらかじめ分類すると見られる。判定することは、正常データに対する損失が引く前提で行うため、横線（オレンジ線）を引き、上が異常データに、下が正常データに判定すると見られる。すると、左下のエリアの存在が問題になる。人気とモデルの損失が不一

致なのである。エリアは左上、左下、右上、右下の順で、それぞれ真陽性、偽陽性、偽陰性、真陰性の数量を表す。右上の偽陰性がほとんどないが、左下の偽陽性が多くあることで、仮説の一条目が証明される。

なお、仮説の二条目は、図9の分布によって証明される。共通パターンに合致する具合によって人気を判断できる場合、点の分布、せめて境界線は右下がりに斜めになるはずである。しかし、図9の分布はそれになっていない。境界線でも段階的な変化が見えない。点の有無の境界線は縦横の二直線に大まかに沿うため、連続的な変化ではなく、離散的な変化だと分かる。つまり、モデルの結果にだけ見れば、正常データのどれが人気であるかも、異常データのどれが一番人気のないかも判断できない。言語道断で言えば、人気作品は質が敷居を跨げれば、どのくらい人気を集めるかは形式上の質との関係が小さいと言えなくもない。

6. 結論

本研究によって、人気作品の量的分析ができた。まず、文字列を単語分散表現に変換することで、ストーリーの展開を人工知能によって分析することを可能にした。そしてオートエンコーダによって、人気小説の共通特徴を抽出することに成功した。訓練されたオートエンコーダは共通パターンの存在か否かによって、人気作品か否かを判定できることを確認できた。そこで、第一種過誤の発生頻度で、仮説一：「ある条件を満たさない作品は、人気小説になる可能性がない。条件を満たす作品は、人気小説になる可能性を有するが、必ず成り上がることが決まらない。」が証明された。なお、モデルの出力の損失に明らかな閾値が存在することで、仮説二：「条件を満たすか否か二択のみあるため、条件の合致する具合で作品を評価することができない。」が証明された。出版などの場面でモデルが活用されるのも期待される。

但し、共通パターンの存在と影響が確認したものの、機械学習の性質上、共通パターンの詳細を分かる方法がない。題材、叙述などの低いレベルにおいて要因それぞれの影響力を分析することは未だ不可能である。というわけで、本研究はモデルの信憑性が確認されながら、説明性の不足を認めざるを得ない。

本研究は従来で個別作品しか取り扱えない物語論を、本研究は量的分析によってその限界を問い詰めた。なんで人気になったかから、人気になったとはどういうことなのかというメタ理論の側面を解き明かした。なお、オートエンコーダ構造の機械学習モデルを実装することで、自動評価における応用の提案も検証された。将来、説明的人工知能の発展とともに、本研究と従来の物語論が連結されれば、自動評価の実用化が期待される。

参考文献

秋元泰介・小方孝（2013）「物語生成システムにおける物語言説機構に向けて―物語言説論と受容理論を導入したシステムの提案--」『認知科学』20：4

キャンベル＝ジョセフ（2015）「千の顔をもつ英雄[新訳版]上」倉田真木訳、早川書房
ジュネット＝ジェラルル（1985）「物語のディスクールー方法論の試み」花輪光訳、水声社

スナイダー＝ブレイク（2010）「SAVE YHE CAT の法則--本当に売れる脚本術」菊池淳子訳、フィルムアート社

Jacob Devlin etc. (2018) BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, arXiv:1810.04805

Kaiming He, etc. (2015) Deep Residual Learning for Image Recognition, CVPR, the Computer Vision Foundation

G. E. Hinton (2006) Reducing the Dimensionality of Data with Neural Networks, SCIENCE vol.313

Ashish Vaswani, Noam Shazeer (2017) Attention is all you need, Advances in Neural Information Processing Systems 30, NeurIPS Proceedings

なろう小説 API マニュアル（2023 年 5 月 13 日アクセス）：

<https://dev.syosetu.com/man/api/#link6>

ソースコードとローデータ

https://github.com/uzuki-sae/About_Syousetsuka_ni_narou.git