

日本語 GPT の蒸留における損失関数の比較

西村稜真¹ 狩野芳伸¹

¹ 静岡大学大学院
{rnishimura,kano}@kanolab.net

概要

本研究では日本語を対象に、蒸留を用いて GPT-2 の小型化のための事前学習を行った。蒸留に際しては、通常の事前学習におけるマスク言語モデルの損失と蒸留の損失の混合率や、蒸留の損失において計算対象とする層など、異なる損失関数を用意して比較した。パープレキシティによる主観評価では、small サイズのベースラインに対しその 68% のサイズの生徒モデルでよりよい値を達成し、文生成結果の主観評価では上回らなかったものの、大きなモデルの性能に匹敵するより小さいモデルへの蒸留の可能性を示した。

1 はじめに

自然言語処理の言語モデルは年々パラメータ数、学習データ量、そのための計算量とコストが連動して増え続けている。もし性能を担保したままより小さいモデルが実現できるならば、有用であることは言をまたない。また、モデルの冗長な部分を探る手掛かりになる。

自然言語モデルの軽量化手法にはいくつかよく知られたものがある。Zafri[1]らは量子化と呼ばれる手法によって BERT[2] の小型化を行い、99% の精度を保ったまま 4 倍軽量のモデルである Q8BERT を構築した。これはモデルパラメータ等の精度を落とすことで軽量化する手法であり、最新の生成モデルにも適用されている [3]。Hinton ら [4] は DNN における蒸留を (はじめて) 行った。蒸留とは、教師モデルの持つ情報を生徒モデルに移す手法である。Sanh ら [5] は教師モデルである BERT の 97% の精度を保ったままパラメータ数を 40% 削減した DistilBERT を作った。

一方で GPT-2[6] をはじめ GPT の蒸留についてはあまり研究されていない。英語版の DistilGPT2 モデルは Hugging Face によって学習されたものが公開さ

れている [7]。RAIL-KD[8] は中間層の出力をランダムに選択し比較する手法であり、RoBERTa に適用している。Li ら [9] は、GPT-2 の蒸留について複数の蒸留手法を検討しており、これには RAIL-KD が含まれる。また同論文では GPT-2 の蒸留の際のパラメータ初期化についても検討している。

日本語版の DistilGPT2 モデルは公開されている [10] ものの学習が収束していないためか、繰り返し同じパターンを生成してしまう状態である。また、GPT の蒸留の際にどのような損失を加えると生徒モデルの生成文の流暢性や一貫性、魅力のいずれに影響が大きいのか検証した論文は知る限りない。

本研究では日本語を対象に、GPT の蒸留について異なる損失を用いて比較し、自動評価および主観評価を行った。

2 関連研究

DistilBERT[5] は教師モデルである BERT の 97% の精度を保ったままパラメータ数を 40% 削減した。DistilBERT は、BERT の通常の事前学習の損失、蒸留の損失、および隠れ層の損失の線型結合を損失関数として用いている [5]。DistilBERT の論文には明示的に式が書かれていないが、該当部分の実装 [11] によると、コサイン類似度の計算には最後の隠れ層の出力のみを用いている。

DistilGPT2 の実装は DistilBERT の手法を適用したとのことなので、これも同様だと思われる。この損失の効果は論文中で検証されているが、中間の隠れ層の出力同士を比較することについての検証はなされていない。

3 提案手法

3.1 生徒モデル

12 層の日本語 GPT-2 モデルである rinna 株式会社の japanese-gpt2-small[12] をベースとして、生徒モデ

ルを次のような方法で初期化した。まず、ベースとなるモデルの層数を半分すなわち 6 層にしたモデルを用意した。その後、japanese-gpt2-small の 1, 2, 5, 8, 11, 12 層目のパラメータを用いてモデルパラメータを初期化した。パラメータ数を表 1 に示す。

3.2 蒸留

蒸留に際しては以下の異なる損失を用い、結果を比較する。

CLM Loss L_{CLM} GPT-2 のオリジナルのマスク言語モデルによる事前学習の損失。予測確率と正解ラベルとのクロスエントロピー誤差をとっている。学習文 S における i 番目の単語を w_i 、生徒モデルの出力する単語 w_i の予測確率を $p_{student}(w_i)$ とすると、式 1 の通りである。

$$L_{CLM} = -\frac{1}{|S|} \sum_{i=1}^{|S|} \log(p_{student}(w_i)) \quad (1)$$

Distillation Loss L_{CE} DistilBERT と同様の蒸留の損失。生徒モデルの出力する予測確率分布と教師モデルの出力する予測確率分布とのクロスエントロピー誤差をとっている。語彙の集合を W 、生徒モデルの出力する単語 w の予測確率を $p_{student}(w)$ 、教師モデルの出力する単語 w の予測確率を $p_{teacher}(w)$ とすると、式 2 の通りである。

$$L_{CE} = -\sum_{w \in W} \{p_{teacher}(w) \cdot \log(p_{student}(w))\} \quad (2)$$

Cosine embedding Loss L_{COS} 隠れ層のコサイン類似度。本研究においては、中間の隠れ層の出力同士を比較し損失に組み込むことを提案する。入力側と出力側の 2 層のパラメータが生徒モデルと教師モデルでそれぞれ一致するように初期化を行っているため、これらの対応が取れていると考えた。すなわち、生徒モデルの第 2 層と教師モデルの第 2 層の出力、および生徒モデルの第 4 層と教師モデルの第 10 層の出力を取り出し、これらのコサイン類似度をとった。生徒モデルの i 層目の隠れ層の出力を $h_{student,i}$ 、教師モデルの j 層目の隠れ層の出力を $h_{teacher,j}$ とすると、実際の計算方法は以下の式 3 である。各項は、損失関数として用いることを想定して 1 から引いた。

$$L_{COS} = \{1 - \cos(h_{student,2}, h_{teacher,2})\} + \{1 - \cos(h_{student,4}, h_{teacher,10})\} \quad (3)$$

表 1 ベースに利用した各モデルのパラメータ数

	層数	パラメータ数
rinna/japanese-gpt2-small	12	134,994,432
rinna/japanese-gpt2-medium	24	368,896,000
生徒モデル	6	92,467,200

表 2 使用した Wikipedia データの統計 (行数)

訓練データ	テストデータ	検証データ
7,364,598	920,562	919,977

3.3 損失関数

学習の際の損失関数には、先に挙げた CLM Loss (L_{CLM})、Distillation Loss (L_{CE})、Cosine embedding Loss (L_{COS}) を組み合わせたものを用いた。

具体的な組み合わせとしては、まず L_{CLM} と L_{CE} との線型結合をいくつか用意した。 L_{CLM} に比べて L_{CE} の値が 2-3 桁ほど小さな値をとったため、適切な学習を行うために L_{CE} に適当な数を掛ける必要があると考えたためである。

また同様の理由で、 L_{CLM} と L_{CE} の相乗平均を取ったものを用意した。最後に、隠れ層の出力を利用した蒸留の効果を確認するため、 L_{CLM} 、 L_{CE} 、 L_{COS} を足し合わせたものを用いた。

なお、GPT-2 の通常の事前学習のロスである L_{CLM} を用いず、 L_{CE} のみを用いて学習を行った際には、学習が収束しなかったため、対象から除外した。

4 実験設定

4.1 データセット

日本語版 Wikipedia の全文をデータセットとして用いた。Wikipedia データは 01-Sep-2023 18:05 版のものを使った。合計の文字数は 11 億文字程度であった。doc タグと空行を削除するなどの前処理を行い、そのうち文字数が 30 以上の行だけを取り出し、訓練データ: テストデータ: 検証データ = 8: 1: 1 となるようランダムに分割した。それぞれのデータ数は表 2 の通りであった。

ハイパーパラメータは表 3 のようにした。バッチサイズについては、教師モデルや実行環境によって表の範囲で変更した。検証時の損失は L_{CLM} のみを用い、この損失において性能の向上が見られなくなった段階で学習を終了させた。

表3 学習条件

パラメータ	値
(生徒モデルの) 層数	6
ヘッド数	12
隠れ層の次元	768
語彙数	32000
最大トークン長	128
最適化手法	Adam
学習率	0.001
バッチサイズ	128-256
最大エポック数	10

表4 モデルのパープレキシティ (低い方が良い値)

	損失関数	教師	PPL
rinna, small	L_{CLM}	—	23.9
rinna, medium	L_{CLM}	—	19.0
proposed model a	L_{CLM}	—	22.0
proposed model b	$L_{CLM} + L_{CE}$	small	21.9
proposed model c	$L_{CLM} + 5 \times L_{CE}$	small	22.4
proposed model d	$L_{CLM} + 10 \times L_{CE}$	small	21.8
proposed model e	$L_{CLM} + 20 \times L_{CE}$	small	22.1
proposed model f	$\sqrt{L_{CLM} \times L_{CE}}$	small	22.4
proposed model g	$L_{CLM} + L_{CE}$	medium	21.5
proposed model h	$L_{CLM} + L_{CE} + L_{COS}$	small	22.4

5 実験結果

5.1 自動評価

パープレキシティを指標としてモデルの性能を測定した。パープレキシティ測定用のデータには、テストデータのうち3000件を取り出して用いた。測定結果は表4の通りである。表中の教師モデルのカラムについて、smallおよびmediumは、それぞれrinnaのsmallおよびmediumを教師モデルとして用いて蒸留したことを意味する。

本研究で学習を行ったモデルはrinnaのsmallの半分の層数であるが、いずれもrinnaのsmallより良いパープレキシティを示し、rinnaのmediumより悪いパープレキシティを示した。

L_{CLM} のみで通常の事前学習を行ったmodel aでは、rinnaのsmallよりパープレキシティが1.9低い値を示した。以下、このmodel aを基準とすると、model bでは0.1、model dでは0.4、model hでは0.5良いパープレキシティの値を示した。model cでは0.4、model eでは0.1、model fでは0.4、model hでは0.4悪いパープレキシティの値を示した。

5.2 人手評価

3名の評価者による人手評価を行った。評価コストのため、先に挙げたモデルのうち同系列のモデル

表5 人手評価のための書き出し

	文章	引用元
A	吾輩は猫である。名前は生命、宇宙、そして万物	夏目漱石 『吾輩は猫である』
B	についての究極の疑問の答えは	huggingface rinnaのデモ [13]
C	イギリスからは紅茶だけでなく、伝統的なティータイムの習慣もオーストラリアに	Wikipedia

と比較して性能の低いmodel cおよびmodel eを除外し、残った8個のモデルについて評価した。

評価に用いるデータは、それぞれのモデルに同じ書き出しから100トークンの文章を生成させて作成した。書き出しとして、表5に示す3種類を用意した。書き出しAは小説的な文章、書き出しBは哲学的な文章、書き出しCは百科事典的な文章を生成させることを意図している。書き出しCはwikipediaに含まれる文章であるが、モデルの学習およびパープレキシティの測定に用いなかったデータを用いた。

文章生成の際のパラメータは以下のようにした。

- temperature: 1.0
- top-k: 500
- top-p: 0.95

生成された文章の例を付録Aに示す。

生成された文章に対して、流暢性(構文)、一貫性(意味・文脈)、魅力(多様性)の三つの指標で主観評価を行った。評価に際しては、異なる二つのモデルの生成文を同時に提示し、それぞれの主観評価指標についてどちらが優れているか選択するペアワイズ形式をとった。選択肢は提示した二つの文と、「全く差がない」の三種類であるが、評価者にはできるだけ「全く差がない」を選ばないように求めた。偏りが生じるのを避けるため、モデルのすべてのペアについて一度ずつ比較させた。すなわち、一つの書き出しにつき28通りの比較がなされた。

主観評価の結果を表6に示す。表の数値はそのモデルの出力を他のモデルの出力と比較した際に、より優れていると判断された割合である。「全く差がない」が1回以上選ばれていた場合、その回数に0.5を乗じた値をより優れていると判断された回数に算入した。

新たに構築したモデルの人手評価の結果は、蒸留を行わなかったmodel aを除いて、全ての項目で教師モデルを下回った。教師モデルを除いた評価の

表6 モデルの人手評価（高い方が良い値）

	流暢性	一貫性	魅力	3項目の平均値
rinna, small	0.56	0.59	0.60	0.58
rinna, medium	0.59	0.54	0.65	0.59
proposed model a	0.52	0.56	0.47	0.52
proposed model b	0.41	0.44	0.30	0.39
proposed model d	0.54	0.49	0.56	0.53
proposed model f	0.54	0.48	0.38	0.47
proposed model g	0.41	0.45	0.52	0.46
proposed model h	0.43	0.44	0.51	0.46

結果は以下の通りである。流暢性が最も高かったのは model d および model f, 次いで model a であった。一貫性が最も高かったのは model a, 次いで model d, model f の順であった。魅力が最も高かったのは model d, 次いで model g, model h の順であった。3項目の平均値では, model d がわずかに model a を上回り, model f がその次に高い値を示した。

6 考察

すべてのモデルが rinna の small の 68% 程度のモデルサイズであるにもかかわらず, より良いパープレキシティを示した。蒸留をせずに通常の事前学習を実行した model a が既に良いパープレキシティを示していることから, これは蒸留の効果ではなく, モデルの初期化方法および Wikipedia による事前学習の効果である。一因として, rinna の事前学習には日本語 Wikipedia および cc100 を用いている [13] が, 本研究においては日本語 Wikipedia のみを用いて蒸留を実行したため, Wikipedia を用いて測定したパープレキシティにおいては本研究のモデルが有利であったと考えられる。

L_{CE} を加えるとわずかにパープレキシティが改善した。複数の条件について調べたところ, L_{CE} に掛ける値によるパープレキシティへの影響はあまり安定していない。特に1倍と10倍の場合には事前学習のみをしたモデルよりもパープレキシティが良かったが, 5倍と20倍の場合には事前学習をしたモデルよりも悪かった。

small からの蒸留より medium からの蒸留の方が性能が良かった。medium の確率分布は small の確率分布よりも正確であると考えられるので, より正確な確率分布がモデルの学習を助けると考えられる。つまり L_{CE} を学習に用いること自体には意味があると考えられる。

L_{CLM} と L_{CE} との相乗平均を取った model f および中間層の出力のコサイン類似度である L_{COS} を利

用した model h では, パープレキシティは向上しなかった。

続いて人手評価の結果について考察する。まず通常の蒸留の効果を確認するために, 蒸留を行わなかった model a と L_{CE} を加えた model b, および $10 \times L_{CE}$ を加えた model d を比較する。model b では全ての項目について model a を大きく下回ったが, model d では流暢性および魅力について model a より高い値を示した。パープレキシティによる評価を考え合わせると, L_{CE} に掛ける値を適切に調整することが重要であると考えられる。

評価項目別に見ると, 流暢性では, model d および model f のみが蒸留を行っていない model a を上回った。他のすべてのモデルと異なりこれらに共通しているのは, L_{CE} の効果が大きく出るように損失関数を再設計したという点であり, L_{CE} がモデルの流暢性に影響を与えると考えられる。

一貫性では, 全ての蒸留モデルが model a を下回った。蒸留は教師モデルの多様な表現を取り込むために, かえって一貫性を損なう可能性がある。

魅力では, 流暢性と異なり, L_{CE} の影響は様々であった。一方で中間層の出力のコサイン類似度である L_{COS} を考えると, これを利用した model h は model b と比べて流暢性, 一貫性では変わらなかったものの, より高い魅力を示した。中間層の出力を利用することで, 教師モデルの表現の多様性を学習することができたと考えられる。

7 おわりに

本研究では生成モデルの蒸留を検討した。パープレキシティを指標とした自動評価では, 通常の前学習に加えて L_{CE} を導入することでモデルの性能を向上させられることを示した。主観評価では, 一貫性以外の項目で蒸留によって主観評価を向上させることのできる損失関数の組み合わせを発見し, 特に L_{CE} に掛ける値を適切に調整することが重要であることを改めて示した。また, 蒸留のために追加された2種類の損失 L_{CE} および L_{COS} について, 前者はモデルの流暢性に, 後者はモデルの魅力に大きな影響を与える可能性が示唆された。

謝辞

本研究は JSPS 科研費 JP22H00804, JP21K18115, JP20K20509, JST AIP 加速課題 JPMJCR22U4, および セコム科学技術財団特定領域研究助成の支援をうけたものです。

参考文献

- [1] Ofir Zafrir, Guy Boudoukh, Peter Izsak, and Moshe Wasserblat. Q8bert: Quantized 8bit bert. In **2019 Fifth Workshop on Energy Efficient Machine Learning and Cognitive Computing - NeurIPS Edition (EMC2-NIPS)**. IEEE, December 2019.
- [2] Jacob Devlin Ming-Wei Chang Kenton and Lee Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. In **Proceedings of naacl-HLT**, Vol. 1, p. 2, 2019.
- [3] Elias Frantar, Saleh Ashkboos, Torsten Hoefler, and Dan Alistarh. Gptq: Accurate post-training quantization for generative pre-trained transformers. **arXiv preprint arXiv:2210.17323**, 2022. <https://doi.org/10.48550/arXiv.2210.17323>.
- [4] Geoffrey Hinton, Oriol Vinyals, and Jeff Dean. Distilling the knowledge in a neural network. **stat**, Vol. 1050, p. 9, 2015.
- [5] Victor Sanh, Lysandre Debut, Julien Chaumond, and Thomas Wolf. Distilbert, a distilled version of bert: smaller, faster, cheaper and lighter. **arXiv preprint arXiv:1910.01108**, 2019. <https://doi.org/10.48550/arXiv.1910.01108>.
- [6] Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, Ilya Sutskever, et al. Language models are unsupervised multitask learners. **OpenAI blog**, Vol. 1, No. 8, p. 9, 2019.
- [7] Hugging Face. distilgpt2 · Hugging Face, (accessed 2024-01) . <https://huggingface.co/distilgpt2>.
- [8] Md Akmal Haidar, Nithin Anchuri, Mehdi Rezagholizadeh, Abbas Ghaddar, Philippe Langlais, and Pascal Poupart. RAIL-KD: RANdom intermediate layer mapping for knowledge distillation. In **Findings of the Association for Computational Linguistics: NAACL 2022**, pp. 1389–1400. Association for Computational Linguistics.
- [9] Tianda Li, Yassir El Mesbahi, Ivan Kobyzev, Ahmad Rashid, Atif Mahmud, Nithin Anchuri, Habib Hajimollahoseini, Yang Liu, and Mehdi Rezagholizadeh. A short study on compressing decoder-based language models. **arXiv preprint arXiv:2110.08460**, 2021. <https://doi.org/10.48550/arXiv.2110.08460>.
- [10] knok/japanese-distilgpt2 · Hugging Face, (accessed 2024-01) . <https://huggingface.co/knok/japanese-distilgpt2>.
- [11] transformers/examples/research_projects/distillation/distiller.py at main · huggingface/transformers · GitHub, (accessed 2024-01) . https://github.com/huggingface/transformers/blob/main/examples/research_projects/distillation/distiller.py.
- [12] rinna 株式会社. rinna/japanese-gpt2-small · Hugging Face, (accessed 2024-01) . <https://huggingface.co/rinna/japanese-gpt2-small>.
- [13] rinna 株式会社. rinna/japanese-gpt2-medium · Hugging Face, (accessed 2024-01) . <https://huggingface.co/rinna/japanese-gpt2-medium>.

A 付録

書き出しおよびモデルごとに、生成された文章の例を表7に示す。紙幅の関係上、生成文の126文字目までを掲載したが、実際にはこの倍程度の長さの文章を生成し、主観評価に用いている。

表7 生成された文章の例

	書き出しA	書き出しB	書き出しC
rinna, small	吾輩は猫である。名前は「ハナコ」。店長から、「アントラクト」と誤って呼ばれ入城した際、店のオーナーに感謝すると同時に、自分の居場所を告げてくれた。アルバイトとして猫を飼っている。頭の数を数えて、全然わからないと言っていた。店長から「ネコ」と呼び掛	生命、宇宙、そして万物についての究極の疑問の答えは、この《私》にあると確信しています。つまり、物の表面が、物全体と同じように見えなくても、物体として見て初めて、その物全体よりも、物はその全体そのものだと思出すことができるのです。また、もう一つ重要な	イギリスからは紅茶だけでなく、伝統的なティータイムの習慣もオーストラリアに伝わった。特にティータイムは、ミルク入りの紅茶が主流であった。イギリスではお茶に含まれるカフェインが血流を良くし、利尿作用や血管拡張・眼精疲労などの副作用の発現が緩やかになる
rinna, medium	吾輩は猫である。名前は未知。ポエム文章力不足で読みにくい文しか書けない不自由な俺の日記(書けたとしても大した内容ではない)。え? 読んでくれたんだけど。読むときと目が腫れちゃうからやめておこう。作者は黒猫ではなく、真逆な自分に似たり寄ったりの男	生命、宇宙、そして万物についての究極の疑問の答えは無敵だ。それは人間の生命の万物の愛よる永遠の哲学、それが宇宙の起源であることは確かだろうが、宇宙は、この星で「生き続ける」ことだ――生命が減り、死が訪れるまで我々が理解できるように、宇宙のことは、宇宙	イギリスからは紅茶だけでなく、伝統的なティータイムの習慣もオーストラリアに浸透している。オーストラリアの紅茶には大きく分けてハーブティー、ジンジャーティー、スパイスティーという3つのタイプがある。一般的にはハーブティーにはルー、スパイスティーにはスパイ
my model a	吾輩は猫である。名前は「仏」の意で、姓は仏教に由来する。釈迦仏は釈迦仏に深い尊敬と尊敬をおかれている。ただし釈迦仏のみについては一説批判がある。一方で、仏教の生物でさえも珍しい。また、仏国土の木として信仰されている。また、サガラの鉢植えや昆虫の幼虫など	生命、宇宙、そして万物についての究極の疑問の答えは、「無限の創造可能的生命が存在する」ということである。つまり、無限の創造物が存在しながら、何が地球以外の諸王国の支配下にあったとしても、生命、宇宙、万物は存在するという主張である。宇宙は完全な創造物であ	イギリスからは紅茶だけでなく、伝統的なティータイムの習慣もオーストラリアに広がり、このティータイムは紅茶の輸出港と輸出区域の入り交じった商業的なプランテーション農業の産物であった。しかし、イギリス本土の貿易路が整備されていなかったため、イギリス本国やア
my model b	吾輩は猫である。名前は兎で、字は五光・龍。眉間・目は僧帽山(現在の法界寺霊園)に由来する。尾は尾、眉は小手。眉は老人で瘤が無い。年齢は不詳。江戸時代は牛馬にまともな馬上槍試合に出ており、記録に残る限り無類の強さを誇る(現在の牛馬一本槍試合などの記録で	生命、宇宙、そして万物についての究極の疑問の答えは、社会史的論争の中で、議論の基礎をなす。この主題はこの主題を含む宇宙論の中心的分野を占めるものであり、たとえば、全ての生命体の生存は核によって決定されると述べられる。ところが、核を含む各生命体が、ある	イギリスからは紅茶だけでなく、伝統的なティータイムの習慣もオーストラリアにはなかった。紅茶はスペアレバーのかわりに茶葉から作った湯により作られた。テーブル上に置かれた茶葉をバターとバターで炒め、緑色の葉だけを使用して焼いた。フランスからは非常に高価な紅
my model d	吾輩は猫である。名前は、猫がそのうろたに抱えたい願いを込めて「うるおい・ぬう」と(ら)いと猫の鳴き声から取られている。家中では「ぬうゃい」、「ぬうゃっし」と呼ばれている。そのため、上が白猫・下が鼠猫である。下が白猫である。そのことから「ぬう」は虎の俗称	生命、宇宙、そして万物についての究極の疑問の答えは、科学の科学的方法論から生まれた。当初は問題を解くだけで科学的な発見もできず、後に原子や分子、あるいはその他の元素に由来する「科学」への批判が展開され、科学がその必要性を消滅させるという解釈もなされた。	イギリスからは紅茶だけでなく、伝統的なティータイムの習慣もオーストラリアに定着した。ミルクティーに紅茶をセットすることも。ただし紅茶はチャクナジャコ(チャクナジャコ)と呼ばれて大衆の飲料とされ、「チャクナジャコ(ca co)」の語はより一般的なものとなり
my model f	吾輩は猫である。名前は、『一、二、六三』の「吾」と「猛」を合わせたもの。お年寄りの者からは「吾」という字に掛けて「参」、お年寄りなら「竹の師」、関四郎なら「竹の師」、兵法三昧なら「兵法三昧堂」や「兵法三昧堂」、「兵法正味堂」などの字に通じる。このよう	生命、宇宙、そして万物についての究極の疑問の答えは、知識世界の形成と認識の重要さの向上、そして思考の世界の形成にある。これを達成するためには、主知や述語などの知的能力、想像力、想像力に関する知識が重要である。科学や技術に対する知識、科学的知識、科学的	イギリスからは紅茶だけでなく、伝統的なティータイムの習慣もオーストラリアに引き継がれている。このティータイムは、第一次世界大戦前後から飲まれていたながら紅茶を飲むというサービスだった。紅茶以外の飲み物も同様に紅茶と同質の紅茶とみなす。また、第二次世界大戦
my model g	吾輩は猫である。名前は羅津から来ている。猫の鱗魚を捕る時、目の通りが、ウサギ耳になった。また、魚が魚を食べないとした時には、貝を美味そうに食べたという。のちに、それを「兎」と呼ぶようになり、さらに十数本の太い鱗魚を捕っていた。この事から、江戸時代には「	生命、宇宙、そして万物についての究極の疑問の答えは、一般に神々が作ったことであると言われる。その根本となるものは神であって「新王国」には悪意を持っておらず、やがて神が減ることとなる。またその自然法則によって「新王国」が存続・発展しないことも間違いではな	イギリスからは紅茶だけでなく、伝統的なティータイムの習慣もオーストラリアにもたらされた。それは英国の紅茶ブランド『ブレンド(rread)』の中心となり、イギリスの茶の生産量を伸ばし、この上質な紅茶がオーストラリアの主な輸出品となった。しかし、1960年
my model h	吾輩は猫である。名前は「猪太郎」である。名前は猫ではなく、「熊五郎」と呼ばれている。年齢は猿楽と、二幕の映画で見られる「ネズミ(幽霊)」を合わせたもの。映画では「まがい物」を表す。『吉宗』での有名な通り、「ごほうび」に出てくる太兵衛は、米粒の卵をかぶせ	生命、宇宙、そして万物についての究極の疑問の答えは、心、思考、思考による精神と、思考と思考に基づく創造の可能性があると述べる。心と思考に基づく精神と思考は、究極の解決策であり、生命では、新たな制約を打ち消すと解釈される。さらに、善、悪と区別される。思考	イギリスからは紅茶だけでなく、伝統的なティータイムの習慣もオーストラリアに受け継がれている。イギリスで紅茶といえば紅茶と紅茶が広く飲まれている。インドでは紅茶専門店がさほど増えておらず、紅茶の消費量は少なく、紅茶の一大輸出国になっていたこともある。また、