

ニューラル言語モデルによる一対多関係知識の記憶と操作

長澤春希¹ Benjamin Heinzerling^{2,1} 乾健太郎^{1,2}

¹ 東北大学 ² 理化学研究所

haruki.nagasawa.s8@dc.tohoku.ac.jp

benjamin.heinzerling@riken.jp inui@ecei.tohoku.ac.jp

概要

ニューラル言語モデル (NLM) による一対多関係知識の記憶と操作の能力について、集合演算を考慮した新しい調査手法を考え、「分類問題」と「生成問題」の2つのアプローチに基づくモデルの振る舞いを調べる研究を進めている。本稿では、調査方法のデザインを論じるとともに、これまで得られた結果として、NLM に一対多関係知識を記憶させることは一定程度可能であるが、集合操作をさせるには別途特別な学習の枠組みが必要であるという示唆が得られたことを報告する。

1 はじめに

近年、自然言語処理分野の多くのタスクで、BERT[1]などの深層ニューラルネットワーク技術を用いたモデルが採用されている。またこれらの深層ニューラル言語モデル (NLM) のパラメータ学習は、事前学習とファインチューニングの2つの過程で行うことが主流となっている。これについて大量のテキストデータを用いた事前学習過程で、文法などの言語知識を獲得しているとされる。

また、NLM は事前学習過程において、エンティティやイベントに関する世界知識の獲得も行なっていることが明らかになっている [2]。このうち、例えば国と首都のような特定の関係のインスタンスを以下では**関係知識**と呼ぶ。ニューラル言語モデルは事前学習の過程でかなり広範な関係知識をテキスト集合から獲得している可能性がある。

こうした背景から NLM を知識ベースと見なしてその能力を調べる研究や NLM と既存知識ベースを統合する研究が広がっている (Language Models as Knowledge Bases) [2, 3, 4, 5, 6]。ただし、こうした研究では、これまでのところ関心の対象は一対一対応の関係知識に留まっており、多対多関係の知識について扱いが簡単でないということ以外明らかにされ

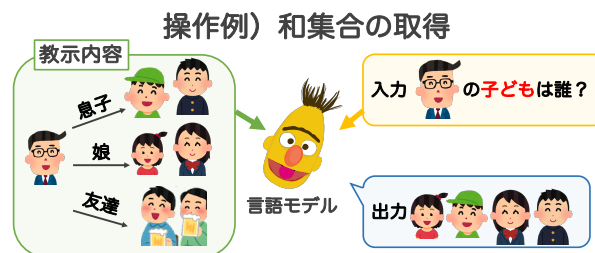


図1 一対多関係知識の記憶と操作のイメージ

ていない [2]。一方、世界に関する知識には、「出演俳優と出演映画」関係のような、多対多関係の知識も多い [3]。従って、NLM を知識ベースと見なすには、NLM が多対多関係知識をどの程度正確に記憶できるか、また記憶した知識の操作 (例えば、多対多関係のインスタンスに属する集合の間の演算) がどの程度可能か、といった能力を精緻に確認する必要がある。そこで本研究では、その第一歩として、一対多関係の知識を人工的に生成したデータセットを構築し、これに対する NLM の記憶と操作の能力を組織的に調べた。

ここで調べたいのは図1に示すような記憶と操作の能力である。「記憶」の能力とは、NLM に〈息子〉関係や〈娘〉関係のような多対多関係のインスタンスを教示した場合に、その教えたインスタンスの情報を NLM に問い合わせて取り出すことができるかである。「操作」の能力としては、例えば図1のように、ある人の〈息子〉の集合と〈娘〉の集合を教示した NLM から、その和集合 (〈子ども〉の集合) を取り出すことができるか、といったように「記憶した集合間の演算を要する問い合わせ」にどの程度答えることができるかを考える。

本研究では、親子関係等の人間関係について関係インスタンスの集合を人工的に合成し、表1に示すような記憶と操作の課題を設計して、NLM の能力を調べた。

データを人工的に合成するのは、統制したデー

表 1 一対多関係知識の集合演算的扱い. モデル内部に特定の知識が格納されている前提のもとで, "期待する出力"に示すような出力が達成されれば, 一対多関係知識に対する集合操作能力が獲得されたと考えられる.

モデルへの教示内容	達成したい操作	期待する出力
・〈子ども〉関係	要素の取得	子どもの名前を個別に出力
・〈子ども〉関係	集合の取得	子どもの名前を全て出力
・〈息子〉, 〈娘〉関係 ・(息子と娘は"子ども"であること)	和集合の取得	(子どもとして,) 息子と娘の名前を全て出力
・〈子ども〉関係	部分集合の取得	名前が3文字の子どもを全て出力
・〈子ども〉関係 ・〈学校 A の生徒〉関係	積集合の取得	学校 A に所属する子どもの名前を全て出力
・〈子ども〉関係 ・〈学校 A の生徒〉関係	集合に対する否定演算	学校 A に所属していない子どもの名前を全て出力
・〈子ども〉関係 ・その子どもの〈子ども(孫)〉関係 ・子どもの子どもが孫であること	集合に対するマルチホップ演算	子どもを介して, 孫の名前を全て出力

タで NLM の能力を調べることがまずは重要だと考えたためである. 実世界のデータを使った extrinsic な評価は, 統制したデータによる intrinsic な評価によって NLM の基本的な能力が明らかになった後に行っていききたい.

2 調査手法

2.1 用語定義

本研究で登場する用語と意味を定義する.

一対多関係知識 1つの主語 (subject) に対し, 特定の関係 (relation) を満足する目的語 (objects) が複数存在するような世界知識のことを指す.

(例: Taro has children named Ichiro, Jiro.)

個別要素 一対多関係知識において, 主語と目的語を一対一対応の形式で結んだ個別の関係知識を指す.

(例: Taro has a child named Ichiro.)

個別教示 一対多関係知識を構成するそれぞれの要素について, 一対一対応の形式で個別に言語モデルに学習させること.

(例: Who is the child of Taro? → Ichiro.)

集合教示 一対多関係知識全体を言語モデルに学習させること.

(例: Who are the children named Taro? → Ichiro, Jiro.)

2.2 調査手法の決定

本研究では, 事前学習済み NLM に対し, 一対多関係知識を明示的に学習させた際の振る舞いを調査する. これについて具体的に「分類問題」と「生成問題」の2つのアプローチを考える. この概念図を図 2 に示す.

具体的な調査方法については, 言語モデルを用いた穴埋め形式と, 生成モデルを用いた回答文生成形式の2つが考えられる. 前者については代表的な先行研究 [2] でも用いられており, モデルの予測分布から最も確率値の高いものをモデルの回答として扱っていた. 本研究はこれに倣い, 確率の高い上位 N 個の目的語を, 一対多関係知識に対するモデルの回答とする. 本研究では, このアプローチ方法を「分類問題」と呼ぶこととする.

また後者については seq-to-seq モデル [7] を用いることで, 複数の目的語を過不足なく生成できるかどうかを考える. 本研究では, このアプローチ方法を「生成問題」と呼ぶこととする.

3 実験

本論文で報告する実験結果については, 主に表 1 の「集合の取得」と「和集合の取得」とする.

3.1 実験データ

本実験では NLM の記憶精度を正確に管理できるようにするため, 学習データを独自に用意した. 具体的には, Wikidata[8] から父親とその子ども (2-4 人) の名前を取得したものを利用した. また使用する人物名について, 次の条件を満たすような調整を行った. (i) 同姓同名の父親 (主語) が存在しない. (ii) 目的語の表層が重複する場合, 区別できるように名前の後ろに id 番号を付与.

また分類問題においては上記の条件を満足しながら, 関係知識データを人工的に作成し, 追加した. 各実験で用いた学習データの具体例を Appendix A の表 3 に掲載する.

なお本研究では, 明示的に学習させた関係インス

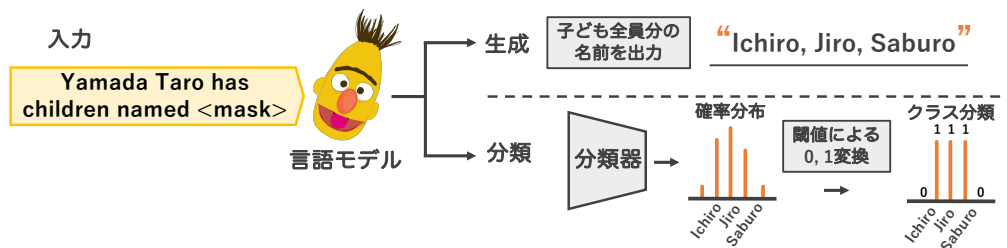


図2 調査手法概念図. 本研究では一対多関係知識について、生成問題と分類問題の2つのアプローチを取る. 生成問題(上段)では複数の目的語を漏れなく生成することを目的とする. 分類問題(下段)では言語モデルの出力を、目的語数次元に変換する分類器に渡す. その後閾値を用いてマルチクラス分類問題の形式へと変換し、複数の目的語を取得する.

タンスの「記憶」と「操作」のみを考えることとする. 従って、学習データとして登場していない関係インスタンスに対する汎化能力などの測定は行わない.

3.2 A: 分類問題

ここでは、先述の「確率分布から上位N個の目的語を取得する」方法を考える. 出力すべき要素を予測確率分布から判別できるかどうかについて、1つ1つの目的語を独立したクラスとして扱うマルチクラス分類問題の形式を考える. 主に一対多関係知識について個別教示を行い、最終的にどのような分布が出力されるかを調査する.

NLMとして、事前学習済み RoBERTa base[9]を使用した. 加えて、このモデルの最終隠れ層の値を目的語数次元のベクトルに変換する線形層を用いた. 推論時にはこの線形層の出力の Softmax を取り、閾値¹⁾より大きいものを1、そうでないものを0に置き換えることで、マルチクラス分類問題の形式に変換した.

3.2.1 実験 A-1: 個別教示によるモデルの振る舞い

本実験では主語と特定の関係からなる入力クエリに対し、目的語を1つ予測する個別教示(例:「Who is the child of Yamada Taro?」に対し「28 (Ichiro のクラス番号)」を予測するよう学習)を全ての目的語について行うことで、個別要素の記憶を図った. この「記憶」の元でNLMの「操作」能力を計測した. 具体的には、目的語集合を問う入力に対し、複数の目的語を過不足なく予測できているかを測った(例:「Who are the children of Yamada Taro?」に対し「28, 29, 30 (Ichiro, Jiro, Saburo のクラス番号)」が予測できているか).

また関係の種類を複数(〈子ども〉, 〈友人〉, 〈同

僚〉など)学習させることで、主語と目的語の共起関係だけでなく、関係の種類にも注目して出力すべき個別要素を選別できるかを検証した. データについては、子どもがN人いる主語に対して、その関係に対する目的語をN人作成するといった形で機械的な追加を行なった.

結果として、関係の種類に注目して正解の目的語を過不足なく予測できることが確認された(図3, 個別要素の記憶精度及び集合の取得精度を示す). 従って主語と関係の種類両方に注目し、いくつの個別要素を出力すべきか、言い換えればどの目的語集合を出力すべきかを正しく判断できていると考えられる. 一方で、図としては掲載していないが、目的語集合の取得が難しい関係の種類も確認された.

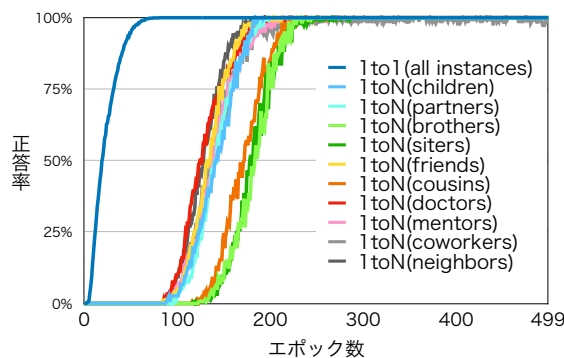


図3 個別要素の記憶精度(1to1)と集合の取得精度(1toN)

3.2.2 実験 A-2: 〈関係の種類〉の和集合の取得

ここでは意味の近い関係の目的語集合を統合する操作、すなわち表1における「和集合の取得」や図1のような出力が可能かを調査する実験を行なった. 具体的には〈息子〉, 〈娘〉関係知識を個別教示した上で、これらを〈子ども〉としてまとめて出力することができるかを調査した. また関係の種類に着目した統合ができているかを確認するため、**子ど**

1) 閾値は Appendix B の式 1 により算出.

も集合に含まれるべきではない〈友人〉関係も同時に学習させた。

息子と娘が"子ども"であることをモデルに学習させる素朴な方法として、一定数の主語をサンプルし、その息子と娘インスタンスについて、**son (daughter) テンプレート** ("Who is the son (daughter) of S?") と **child テンプレート** ("Who is the child of S?") の両方を利用した学習を行った。

結果として、children を尋ねるクエリに対し、息子と娘の両方を取得することが可能であることが確認された(表 2)。事前学習済みモデルであれば、child テンプレートを用いずとも、半数以上のインスタンスについて、"子ども"として息子と娘の両方を出力できることが明らかとなった。また child テンプレートの使用率を大きくすることで、和集合の取得率も向上することが確認された。

なお今回検証したのは上記 1 パターンのみであるが、和集合の取得が可能かどうかについては、関係の種類 of 包含関係を事前学習済みモデルがどれ程適切に扱うことができるかで結果が大きく左右されることが考えられた。

表 2 和集合の最大取得率

child テンプレート 使用率 (%)	children クエリの 最大正答率 (%)
0	65
10	77
30	96
50	100

3.3 B: 生成問題

一対多関係知識を扱うもう 1 つのアプローチとして、事前学習済み BART base[10] を用いた回答テキストの生成を考え、集合教示と個別教示の両方を行った。

3.3.1 実験 B-1: 個別教示によるモデルの振る舞い

ここでは個別教示による学習を行い、目的語を統合して出力することが可能かを調査した。個別教示については"S has a child named <mask>"に対し、子どもの名前を 1 つ生成させる学習を行った。

結果として、個別要素は全て正しく記憶できるがそれらを統合して出力することは難しいということが明らかになった。具体的には、個別教示のみを行う設定であると複数目的語を出力させるクエリ (S has children named <mask>) に対しても、目的語を 1 つしか出力できないことが確認された。

3.3.2 実験 B-2: 個別教示と集合教示

実験 B-1 (3.3.2 節) の個別教示に加え、一定数の主語に対し、集合教示を行った場合のモデルの振る舞いを調査した。集合教示としては、「S has children named <mask>」の入力文に対し、子ども全員分の名前をカンマつなぎで出力させる形式を取った。

この場合でも実験 B-1 (3.3.1 節) で確認されたように、全ての個別要素の記憶は問題なく達成できた。しかしながら、明示的に集合教示を行ったインスタンスに対してのみ目的語を複数出力する結果となり、集合を取得するという汎化能力の獲得には至らなかった。

従って「記憶している関係知識に対し、ある種の集成的な操作をすること」を生成問題の枠組みで達成するためには、他に特別な学習方法を用意する必要があることが示唆された。

4 おわりに

本研究では一対多関係知識に対するニューラル言語モデルの振る舞いを、**分類問題**と**生成問題**の 2 つの観点から調査した。本論文で調査した限りではどちらの手法についても、一対多関係知識を言語モデルで扱うことは現状難しいという結論に至った。

分類問題では特定のテンプレートを用いて個別要素を記憶させることで、一対多関係知識をマルチクラス分類問題の枠組みで取り扱える可能性があることを示した。一方で複数の目的語に対する集合演算的な操作については、率直には達成されないことが明らかになった。

生成問題については個別要素の記憶は可能だが、複数個の目的語を出力することは達成されなかった。記憶された要素に対する集合演算的な操作を学習させるための、何らかの枠組みが必要となることが示された。

以上 2 点の結果を踏まえ、保存された個別要素に対する集合演算能力を獲得できれば、一対多関係知識に対しても十分に対応できると考えられる。これについて、集合の扱いとして Deep Sets[11], Set Transformer[12], 学習手法として Meta Learning[13, 14]などを参考に、上記能力の獲得が実現されるかどうかを引き続き調査していく所存である。

謝辞

本研究は JSPS 科研費 21K17814 および JST, CREST, JPMJCR20D2 の助成を受けたものです。

参考文献

- [1] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: pre-training of deep bidirectional transformers for language understanding. In Jill Burstein, Christy Doran, and Thamar Solorio, editors, **Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2019, Minneapolis, MN, USA, June 2-7, 2019, Volume 1 (Long and Short Papers)**, pp. 4171–4186. Association for Computational Linguistics, 2019.
- [2] Fabio Petroni, Tim Rocktäschel, Sebastian Riedel, Patrick S. H. Lewis, Anton Bakhtin, Yuxiang Wu, and Alexander H. Miller. Language models as knowledge bases? In Kentaro Inui, Jing Jiang, Vincent Ng, and Xiaojun Wan, editors, **Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing, EMNLP-IJCNLP 2019, Hong Kong, China, November 3-7, 2019**, pp. 2463–2473. Association for Computational Linguistics, 2019.
- [3] Benjamin Heinzerling and Kentaro Inui. Language models as knowledge bases: On entity representations, storage capacity, and paraphrased queries. In Paola Merlo, Jörg Tiedemann, and Reut Tsarfaty, editors, **Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume, EACL 2021, Online, April 19 - 23, 2021**, pp. 1772–1791. Association for Computational Linguistics, 2021.
- [4] Gerhard Weikum, Xin Luna Dong, Simon Razniewski, and Fabian M. Suchanek. Machine knowledge: Creation and curation of comprehensive knowledge bases. **Found. Trends Databases**, Vol. 10, No. 2-4, pp. 108–490, 2021.
- [5] Tara Safavi and Danai Koutra. Relational world knowledge representation in contextual language models: A review. **CoRR**, Vol. abs/2104.05837, , 2021.
- [6] Simon Razniewski, Andrew Yates, Nora Kassner, and Gerhard Weikum. Language models as or for knowledge bases. **CoRR**, Vol. abs/2110.04888, , 2021.
- [7] Ilya Sutskever, Oriol Vinyals, and Quoc V. Le. Sequence to sequence learning with neural networks. In Zoubin Ghahramani, Max Welling, Corinna Cortes, Neil D. Lawrence, and Kilian Q. Weinberger, editors, **Advances in Neural Information Processing Systems 27: Annual Conference on Neural Information Processing Systems 2014, December 8-13 2014, Montreal, Quebec, Canada**, pp. 3104–3112, 2014.
- [8] Denny Vrandečić and Markus Krötzsch. Wikidata: a free collaborative knowledgebase. **Communications of the ACM**, Vol. 57, No. 10, pp. 78–85, 2014.
- [9] Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. Roberta: A robustly optimized BERT pretraining approach. **CoRR**, Vol. abs/1907.11692, , 2019.
- [10] Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Veselin Stoyanov, and Luke Zettlemoyer. BART: denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. In Dan Jurafsky, Joyce Chai, Natalie Schluter, and Joel R. Tetraault, editors, **Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, ACL 2020, Online, July 5-10, 2020**, pp. 7871–7880. Association for Computational Linguistics, 2020.
- [11] Manzil Zaheer, Satwik Kottur, Siamak Ravanbakhsh, Barnabás Póczos, Ruslan Salakhutdinov, and Alexander J. Smola. Deep sets. **CoRR**, Vol. abs/1703.06114, , 2017.
- [12] Juho Lee, Yoonho Lee, Jungtaek Kim, Adam R. Kiorek, Seungjin Choi, and Yee Whye Teh. Set transformer: A framework for attention-based permutation-invariant neural networks. In Kamalika Chaudhuri and Ruslan Salakhutdinov, editors, **Proceedings of the 36th International Conference on Machine Learning, ICML 2019, 9-15 June 2019, Long Beach, California, USA**, Vol. 97 of **Proceedings of Machine Learning Research**, pp. 3744–3753. PMLR, 2019.
- [13] Chelsea Finn, Pieter Abbeel, and Sergey Levine. Model-agnostic meta-learning for fast adaptation of deep networks. In Doina Precup and Yee Whye Teh, editors, **Proceedings of the 34th International Conference on Machine Learning, ICML 2017, Sydney, NSW, Australia, 6-11 August 2017**, Vol. 70 of **Proceedings of Machine Learning Research**, pp. 1126–1135. PMLR, 2017.
- [14] Pride Kavumba, Benjamin Heinzerling, Ana Brassard, and Kentaro Inui. Learning to learn to be right for the right reasons. In Kristina Toutanova, Anna Rumshisky, Luke Zettlemoyer, Dilek Hakkani-Tür, Iz Beltagy, Steven Bethard, Ryan Cotterell, Tanmoy Chakraborty, and Yichao Zhou, editors, **Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2021, Online, June 6-11, 2021**, pp. 3890–3898. Association for Computational Linguistics, 2021.

A 各実験における学習データの具体例

ここに、4つの実験で用いた学習データの例と、そのデータを用いて学習したモデルに対して期待する出力例を示す。本研究では、表中に示す学習データによって「個別要素の記憶」と「集合演算能力」の獲得を図った。そして、その学習の元で「期待するモデルの挙動例」列に示すような、複数の目的語を出力する能力、すなわち一対多関係知識を適切に扱う演算能力が実際に獲得されているかを調査した。

表3 各実験における学習データ例

実験	学習データ例 (入力 → 出力)	期待するモデルの挙動例 (入力 → 出力)
分類問題 A1	<ul style="list-style-type: none"> • Who is the child of Yamada Taro? → 28 (Taro の子ども 1 のクラス番号) • Who is the child of Yamada Taro? → 29 (Taro の子ども 2 のクラス番号) • Who is the child of Yamada Taro? → 30 (Taro の子ども 3 のクラス番号) 	<ul style="list-style-type: none"> • Who are the children of Yamada Taro? → 28, 29, 30
分類問題 A2	<ul style="list-style-type: none"> • Who is the son of Yamada Taro? → 11 • Who is the son of Yamada Taro? → 12 • Who is the daughter of Yamada Taro? → 24 • Who is the daughter of Yamada Taro? → 25 • Who is the friend of Yamada Taro? → 87 • Who is the friend of Yamada Taro? → 88 (• child テンプレートの使用例: Who is the child of Yamada Taro? → 11) 	<ul style="list-style-type: none"> • Who are the children of Yamada Taro? → 11, 12, 24, 25
生成問題 B1	<ul style="list-style-type: none"> • Yamada Taro has a child named <mask> → Ichiro • Yamada Taro has a child named <mask> → Jiro 	<ul style="list-style-type: none"> • Yamada Taro has children named <mask> → Ichiro, Jiro
生成問題 B2	<ul style="list-style-type: none"> • Yamada Taro has a child named <mask> → Ichiro • Yamada Taro has a child named <mask> → Jiro • Yamada Taro has children named <mask> → Ichiro, Jiro • Mori Akira has a child named <mask> → Kenji • Mori Akira has a child named <mask> → Kyoko • Mori Akira has a child named <mask> → Tomomi 	<ul style="list-style-type: none"> • Yamada Taro has children named <mask> → Ichiro, Jiro • Mori Akira has children named <mask> → Kenji, Kyoko, Tomomi

B 分類問題の正答率および閾値算出について

図3では個別教示による記憶精度と、それを元にした集合取得の正答率を共に示している。前者の記憶精度については、"Who **is** the {relation} of S?"の入力クエリに対して出力する確率分布において、最も確率値の高い目的語が、正解のN個の目的語のいずれかである場合に正解とした。

後者の正答率については、まず"Who **are** the {relation} of S?"を入力クエリとし、モデルの出力分布を得た。Softmax 関数を適用し、各クラスに割り当てられた値を確率として解釈できるよう変換した。この分布を元に閾値を設定し、閾値を上回るクラスを1、そうでないものを0としてマルチクラス分類問題の形式に変換した(図2の下段参照)。

なお閾値の設定については、チェビシェフの不等式を利用した。具体的には1つのクラスを x 、そのクラスに割り当てられた確率を $p(x)$ 、および Softmax 関数を適用した分布の平均値、標準偏差をそれぞれ \bar{x}, s とした時に式(1)を満たすものを外れ値、すなわち取得する目的語として算出した。

$$p(x) > \bar{x} + s \quad (1)$$