

# LLM 推定ラベルと弱教師あり学習による 反復的アノテーション更新

浅野輝<sup>1,2</sup> 小津野将<sup>3</sup> 馬場雪乃<sup>1</sup>

<sup>1</sup> 東京大学 <sup>2</sup>RIKEN AIP <sup>3</sup>OMRON SINIC X  
asano-hikaru19, yukino-baba@g.ecc.u-tokyo.ac.jp,  
tadashi.kozuno@sinicx.com

## 概要

本研究では、大規模言語モデル (LLM) による初期アノテーションと、弱教師あり学習の一種である Robust Unlabeled-Unlabeled Learning を組み合わせた反復的な学習フレームワークを提案する。従来、専門領域や大規模データセットにおけるテキスト分類では、アノテーションコストとラベルノイズが精度向上の大きな障壁となっていた。本手法ではまず LLM を用いて未ラベルデータに対し疑似ラベルを一括付与し、正例比率が相対的に高いコーパスと低いコーパスを疑似正例・負例として構築する。次に Robust UU Learning を適用することで、LLM が付与したノイズを含むラベルにも頑健な分類器を学習し、その分類器で再度データをラベリングする手順を複数回繰り返す。これにより、初期アノテーションの誤りを段階的に削減し、高精度な分類器を獲得できることを実験的に確認した。また、本手法により、GPT-4o などの言語モデルの性能を上回る成果を得られることも示した。

## 1 はじめに

テキスト分類は、感情分析や文書カテゴリの分類といった従来のタスクから、近年は大規模言語モデル (Large Language Model: LLM) の RLHF や DPO などの学習手法や評価に至るまで、幅広い分野で重要な役割を担っている。テキスト分類の有用性が広く認識されている一方で、その学習には大量のラベル付きデータが必要となる。アノテーション作業には高い人的コストが伴い、とりわけ医療や法律など専門性の高いテキストの場合、このコストはさらに深刻な問題となる。アノテーションのコスト削減を目的として、近年は LLM を活用したアノテーション

の自動化が活発に検討されているものの、実際の性能には依然として限界があるとの報告が多い。こうした LLM 由来のアノテーション・ノイズを低減するため、複数エージェントの相互評価による進化的アプローチが提案されているが、ノイズを完全に排除することはなお困難とされている。

このようなノイズを含むアノテーションに対処する方法として、近年では弱教師あり学習が注目されている。弱教師あり学習は、ラベル情報が完全ではない場合や、一部のデータのみがラベル付きであるような不完全な状況でも、モデルを頑健に学習させるための枠組みである。弱教師あり学習の一手法として提案されているのが Unlabeled-Unlabeled (UU) Learning (1) である。UU Learning では、未ラベルデータから 2 種類のコーパス、「疑似正例コーパス (D)」と「疑似負例コーパス (D')」を用意し、「D のほうが D' よりも真のポジティブサンプルの含有率が高い」という緩やかな前提が満たされれば、ラベルなしデータのみを用いても分類器の学習が可能となる。これにより、ラベルの完全性や正確性を厳密に担保できない現場においても、一定の精度を確保しやすい利点をもつ。

そこで本研究では、LLM の推論能力と UU Learning を組み合わせることにより、大規模データのラベリングコストとラベルノイズの問題を同時に軽減する新たな反復学習フレームワークを提案する。まず、LLM を用いて未ラベルのコーパスに対し初期的なラベル付けを行い、LLM が「正」と判断したサンプルを疑似正例コーパス D、LLM が「負」と判断したサンプルを疑似負例コーパス D' として作成し、UU Learning を適用する。この手順によって、元の LLM が有する分類性能を上回る精度をもつ分類器を学習可能となる。

さらに、得られた分類器を用いて元のコーパスを再度ラベリングし、更新された疑似正例コーパス  $\mathcal{D}$  と疑似負例コーパス  $\mathcal{D}'$  を再び UU Learning のパイプラインに投入する。このように反復的にラベルを精錬することで、初期段階で LLM が付与したラベルノイズを段階的に削減し、 $\mathcal{D}$  と  $\mathcal{D}'$  がより理想的な分布へと近づいていく。その結果、学習を重ねるにつれ、分類器の精度向上が期待できる。

本研究では、上記フレームワークによって LLM と UU Learning を組み合わせた反復的学習手法を実装し、その性能向上を実験的に検証する。さらに、その最終的な性能が、従来の自己進化的な LLM 活用手法を上回ることを示し、提案手法の有効性を明らかにした。

## 2 準備

### 2.1 二値分類と教師あり学習

多くの実世界のタスクでは、入力サンプル  $x \in \mathbb{R}^d$  が正例 (+1) または負例 (-1) のいずれかに属する二値分類問題が現れる。これらのサンプルは未知の同時分布  $p(x, y)$  に従って生成されると仮定できる。正例の事前確率を  $\pi_+ = p(y = +1)$  とし、 $p_p(x) = p(x | y = +1)$ 、 $p_n(x) = p(x | y = -1)$  と定義すると、周辺分布は

$$p(x) = \pi_+ p_p(x) + (1 - \pi_+) p_n(x)$$

で表される。

このとき、分類器  $g: \mathbb{R}^d \rightarrow \mathbb{R}$  はサンプル  $x$  に実数スコアを割り当て、その符号 (正負) によって予測ラベルを決定する関数であり、たとえばニューラルネットワークなどが典型的な例である。損失関数  $\ell: \mathbb{R} \times \{+1, -1\} \rightarrow [0, \infty)$  は予測結果と真のラベルとの不一致度合いを測る指標であり、ラベルが正である時の損失  $\ell(g(X), +1)$  を  $\ell^+(g(x))$ 、ラベルが負である時の損失  $\ell(g(X), -1)$  を  $\ell^-(g(x))$  とすると、真のリスクは以下のように表される。

$$\begin{aligned} R(g) &= \mathbb{E}_{(X,Y) \sim p} [\ell(g(X), Y)] \\ &= \pi_+ \mathbb{E}_{X \sim p_p} [\ell^+(g(X))] + (1 - \pi_+) \mathbb{E}_{X \sim p_n} [\ell^-(g(X))] \end{aligned} \quad (1)$$

教師あり学習では正例サンプルと負例サンプルがそれぞれ  $\mathcal{D}_p = \{x_1, \dots, x_n\} \sim p_p(x)$ 、 $\mathcal{D}_n = \{x'_1, \dots, x'_n\} \sim p_n(x)$  として与えられている状況を想定し、これを用いて、式 (1) の経験的リスクを最小化することで、分類器  $g$  の学習を行う。

多くの実問題では、十分な数の正例・負例があれば、高い精度の分類器を学習できることが知られている。一方で、膨大な正例・負例をアノテーションで得るのは容易ではなく、特に専門知識が必要な領域ではアノテーションコストは実用上、大きな問題となる。

### 2.2 Unlabeled-Unlabeled (UU) Learning

UU Learning は、すべてのサンプルに対してラベル付けされた正例・負例のデータセットが直接得られない状況でも、クラス事前分布が異なる 2 つの未ラベルデータセットを用意できれば、分類器を学習できるという枠組みである。

具体的には、2 つの未ラベルデータ  $\mathcal{D} = \{x_1, \dots, x_n\}$ 、 $\mathcal{D}' = \{x'_1, \dots, x'_n\}$  が、それぞれ異なる正例比率  $\theta \neq \theta'$  をもつ混合分布から生成されたと仮定する。すなわち

$$\begin{aligned} p_\theta(x) &= \theta p_p(x) + (1 - \theta) p_n(x) \\ p_{\theta'}(x) &= \theta' p_p(x) + (1 - \theta') p_n(x). \end{aligned}$$

UU learning では  $\theta > \theta'$  が成り立つ時、UU learning の学習が可能になり、正例比率が大きい  $\mathcal{D}$  を疑似正例コーパス、正例比率の小さい  $\mathcal{D}'$  を疑似負例コーパスと呼ぶ。UU learning では疑似正例データセット  $\mathcal{X}$  と疑似負例データセット  $\mathcal{X}'$  から式 (1) で示される二値分類におけるリスクに対して、不偏推定量を構築することが可能であり、UU learning のリスクは

$$\begin{aligned} R_{uu}(g) &= \mathbb{E}_{X \sim p_\theta} [a \ell^+(g(X)) - b \ell^-(g(X))] \\ &\quad + \mathbb{E}_{X' \sim p_{\theta'}} [-c \ell^+(g(X')) + d \ell^-(g(X'))] \end{aligned} \quad (2)$$

と表すことができる。ここで、 $a, b, c, d$  は  $\pi_+, \theta, \theta'$  から求められる重みであり  $a = \frac{(1-\theta')\pi_p}{\theta-\theta'}$ 、 $b = \frac{\theta'(1-\pi_p)}{\theta-\theta'}$ 、 $c = \frac{(1-\theta)\pi_p}{\theta-\theta'}$ 、 $d = \frac{\theta(1-\pi_p)}{\theta-\theta'}$  として定義される。

### 2.3 Robust UU Learning

式 (2) により、個別のサンプル  $x$  に対してラベル付けがされていなくても分類器の学習が可能になる。一方で、元の二値分類のリスク (1) と UU learning におけるリスク (2) を見比べてみると、元のリスクは非負であるのに対して、(2) の損失関数は、 $-b \ell^-(g(X))$  や  $-c \ell^+(g(X'))$  といった「負の項」を含むため、モデルによっては損失が負の方向に大きく振れて訓練データに対して過学習を引き起こすケースがあることが指摘されている。

そこで、一般化 Leaky ReLU 関数  $f$  を導入し、負のリスクを過剰に低減しないように調整する手法が

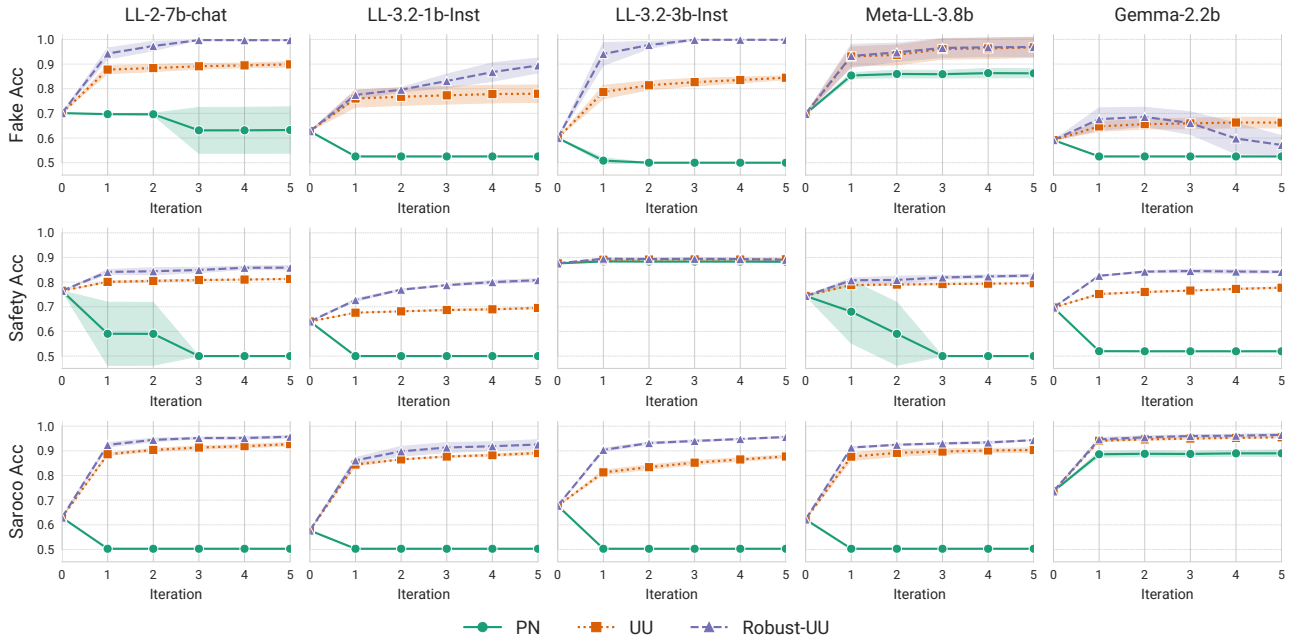


図1 PN、UU、Robust-UU Learning の比較

Robust UU Learning (2) である。具体的には、

$$R_{ruu}(g) = f(a\mathbb{E}_{X \sim p_\theta}[\ell^+(g(X))] - c\mathbb{E}_{X' \sim p_{\theta'}}[\ell^+(g(X'))]) \\ + f(d\mathbb{E}_{X' \sim p_{\theta'}}[\ell^-(g(X'))] - b\mathbb{E}_{X \sim p_\theta}[\ell^-(g(X))]) \quad (3)$$

のように、損失関数の各項を正規化することで、過学習を緩和する仕組みを導入する。ここで、 $f(x)$  は  $f(x) = \mathbb{I}_{\{x>0\}}x + \mathbb{I}_{\{x<0\}}\lambda x, (\lambda < 0)$  のように定義され、負リスクが学習を誤った方向に進めないよう調整する役割を果たす。

### 3 手法

まず初期段階 (Iteration 0) では、LLM を用いて未ラベルコーパス  $\mathcal{U}$  に対し、一括で正例 (+1) または負例 (-1) のラベルを付与する。これにより、LLM が「正」と判断したサンプル集合を疑似正例コーパス  $\mathcal{D}$ 、LLM が「負」と判断したサンプル集合を疑似負例コーパス  $\mathcal{D}'$  として構成する。

得られた  $\mathcal{D}$  と  $\mathcal{D}'$  を用いて、式 (3) で示した Robust UU Learning を実行し、分類器  $g$  を学習する。

Iteration 1 以降は、学習済みの分類器を用いて再度未ラベルコーパス  $\mathcal{U}$  をラベリングし、 $\mathcal{D}$  と  $\mathcal{D}'$  を更新する。この更新されたコーパスを用いて再び Robust UU Learning により新たな分類器  $g$  を学習し、同様のラベル更新と学習のステップを所定の Iteration 回数まで繰り返す。

### 4 実験

本研究では、フェイクであるかどうかを判断する **Fake News Detection Datasets**、低リソース言語であるルーマニア語ニュースの記事が「風刺的内容か否か」を判定する **Saroco**、RLHF 用のデータセットであり、質問に対する回答が安全であるかどうかを判定する **PKU-SafeRLHF** の 3 種類のデータセットを用いて提案手法の評価を行った。

手法 (3) における Iteration 0 における疑似ラベルの作成には、Llama-2-7b-chat-hf、Meta-Llama-3-8B、Llama-3.2-1B、Llama-3.2-3B、gemma-2-2b-it の 5 つのモデルを利用した。学習の際には分類のためにデータセットの説明や正負それぞれの例をプロンプトとして与え、モデルには正負の出力を行わせ、疑似正例コーパス  $\mathcal{D}$  と疑似負例コーパス  $\mathcal{D}'$  を作成した。

分類器の学習では教師あり学習 (PN learning)、UU learning、Robust-UU learning の 3 種類のアルゴリズムを用いた。なお、PN learning の学習の際は疑似ラベルを学習における正負としてそのまま使用して学習を行なった。分類器としては Llama-3.2-1B の出力層に Affine layer を追加して、元の言語モデルの出力を 1 次元に変化する層を追加したモデルを使用して分類を行った。その他の詳細については [A](#) を参照されたい。

UU、Robust-UU learning において必要な正例比率  $\theta$  と  $\theta'$  は訓練データに対してすべてラベルが振られ

ているとして、損失関数を構築し計算を行った。

それぞれの学習の評価としてはテストデータに対する正解率 (Accuracy) を採用した。

## 4.1 実験結果

図 1 に実験の結果を示す。図 1 からわかるように、通常の教師あり学習を使用した場合、LLM の初期アノテーションに誤分類が含まれるため、Iteration が増加するに従って、性能が低下していることがわかる。それに対して、UU、また、Robust-UU Learning を使用した場合、ほとんどのケースにおいて段階的な性能の向上が見られ、また、Robust-UU Learning の方が UU learning に比べて、Iteration の初期から後期に渡ってより高い性能を示していることがわかる。

## 4.2 Ablation Study

前小節の実験では、訓練データ全体に真のラベルが付与されていると仮定し、正例比率  $\theta, \theta'$  を計算したのである。本小節では現実的な設定として、訓練データのうち 50 件または 100 件のみ真のラベルを付与し、これを用いて  $\theta, \theta'$  を計算し、Robust-UU learning を実行した。実験では、3 つの異なるシードを用いて、訓練データからランダムに 50 件または 100 件を選び、ラベルを付与した。図 2 には、50 件、100 件、そして全データにラベルを付けた場合の結果を比較したものを示した。なお、コストを抑えるため、初期のアノテーションモデルとして Llama-2-7b-chat-hf のみを用いて実験を行った。

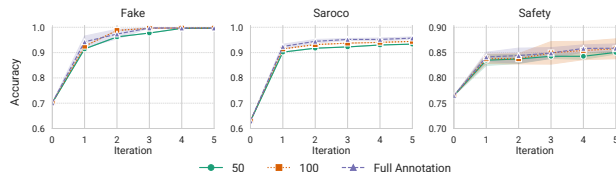


図 2 部分ラベリングによる学習結果

図 2 からわかるように、訓練データに対して 50 件、また、100 件のみという極めて少量のアノテーションをした場合でも、すべてのデータに対してアノテーションをした際の性能と近い性能を示すことが確認され、Robust-UU learning を使用することによって、ラベル付けコストを大幅に削減しつつ、高精度な分類器を構築できることが示された。

## 4.3 難解なデータセットに対する結果

提案手法の有効性を検証するため、Corona (SNS の感情分析)、Green Patent (Green プラスチック関連

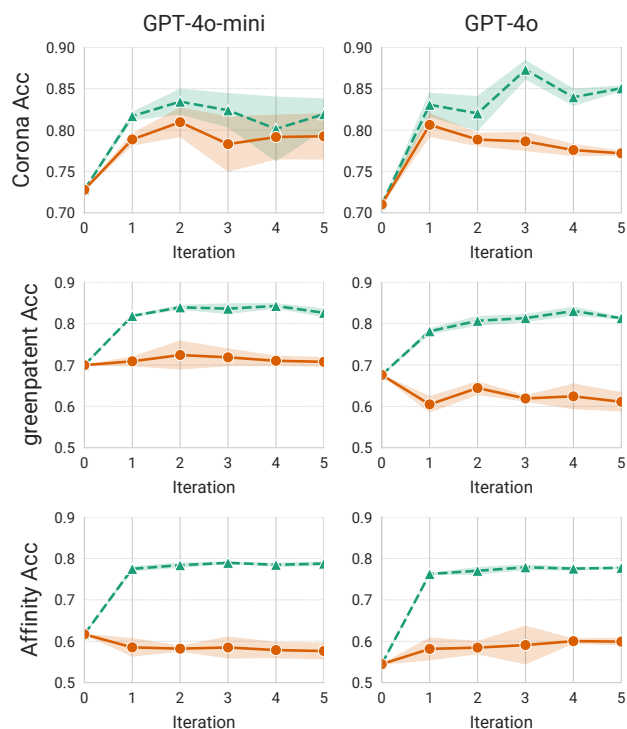


図 3 難解なデータセットに対する実験結果

特許かを分類)、Affinity (分子構造から Covid-19 親和性を予測) に対して追加実験を行った。

アノテーションには GPT-4o-mini と GPT-4o を用い、各データセットで正例・負例 5 個ずつを例示し、モデルはサンプルから抽出した情報・分類理由・ラベルを生成した。また、ベースラインとして GPT-4o-mini の自己修正アルゴリズムを用いた。回答エージェントが出力した情報・理由・ラベルを評価エージェントが評価し、修正を繰り返すことで精度向上を図った。

図 3 からわかるように、提案手法では、GPT-4o などの LLM を活用しても性能向上が難しいデータセットに対して、イテレーション数の増加による性能向上が確認され、専門性の高いテキストへの応用可能性が示唆された。

## 5 おわりに

本研究では LLM によるアノテーションと、Robust Unlabeled-Unlabeled Learning を組み合わせた反復的な学習フレームワークを提案し、GPT-4o などの言語モデルでも分類が難しいタスクに対しても高い性能を示す分類器の学習を実現した。今後は、本研究における提案手法を AI for Science や計算社会科学などへの応用、また、未ラベルのデータセットからの LLM のアラインメントなどへの応用が期待される。

## 謝辞

本研究は、JST ムーンショット型研究開発事業 JPMJMS2236-8 の支援を受けたものです。

## 参考文献

- [1]Nan Lu, Gang Niu, Aditya K. Menon, and Masashi Sugiyama. On the minimal supervision for training any binary classifier from only unlabeled data. In **International Conference on Learning Representations**, 2019.
- [2]Nan Lu, Tianyi Zhang, Gang Niu, and Masashi Sugiyama. Mitigating overfitting in supervised classification from two unlabeled datasets: A consistent risk correction approach. In Silvia Chiappa and Roberto Calandra, editors, **Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics**, Vol. 108 of **Proceedings of Machine Learning Research**, pp. 1115–1125. PMLR, 26–28 Aug 2020.

## A 実験設定の詳細

また、学習の効率化のためにモデルのパラメータを 4b に量子化し、さらに、Lora tuning を採用した。学習においてエポック数はすべて 3 とし、それぞれの Iteration における擬似正例コーパス  $\mathcal{D}$ 、擬似負例コーパス  $\mathcal{D}'$  を使用して各 Iteration におけるモデルの学習を行なった。各 Iteration 終了時には、それぞれのエポック終了後に計算した検証データに対する損失が最も小さいモデルを使用して、再度擬似正例コーパス  $\mathcal{D}$ 、擬似負例コーパス  $\mathcal{D}'$  の作成を行い、また、テストデータに対してラベルの予測も行った。