

入力補完方式を通じた修理部品推薦システムの高精度化：
大規模言語モデルによるテキスト分類モデルのエンハンス
Augmenting Repair Recommendation through Input Completion Method:
Enhancing Text Classification Models with Large Language Models

高平 寛之[†] 水川 秀[†] 飯島 光一郎[†]
Hiroyuki Takahira Hideki Mizukawa Koichiro Iijima

1. はじめに

近年、産業機器の保守事業者では、IoTを背景に故障発生による運転停止時間とメンテナンスコストの削減が求められている。また、少子高齢化による熟練保守員の不足から作業品質のばらつきが問題となっている[1]。このような背景のもと、保守員が記載した症状テキストに基づき、修理に必要な部品番号の予測確率を出力するテキスト分類モデルを用いた修理部品推薦システムが提案されている[2]。保守員はこれら予測確率を参考に修理部品の候補を選定し、効率的な修理対応を実現している。

一方で、予測確率は保守員が記載するテキストの品質に大きく左右され、症状や部位が十分に記載されていない場合は候補部品が絞り込めない増えてしまう。そこで、保守員が高品質なテキストを容易に記載できる仕組みが必要となる。

本研究では、大規模言語モデル(Large Language Model: LLM)を活用し、低品質なテキストから複数の高品質テキスト候補を生成して保守員に提示し、選択させる入力補完方式を提案する。「高品質なテキスト」を「テキスト分類モデルの予測確率が高く、保守員の意図を反映したテキスト」と定義し、以下の2点を工夫した。

- (1) テキスト分類モデルの予測確率が高いテキストのみをLLMに学習させる。
- (2) 修理部品ごとにテキストを生成するようガイドすることで保守員の意図を反映したテキストが候補に含まれるようにする。

提案方式の検証のため、高品質なテキストと低品質テキストを準備し、低品質なテキストから高品質なテキストに補完できるか検証するデータセットを準備した。このデータセットを用いて評価した結果、高品質テキストへの補完成功率は33%に達し、修理部品推薦システムの実質的な正解率を47.8%から65.2%に向上できる見込みを得た。

2. 目標と課題

2.1 目標

本研究では、LLMを用いて低品質なテキストから複数の高品質テキスト候補を生成し、保守員に提示・選択させる入力補完方式を目指す。図1は修理部品推薦システムでの利用イメージである。保守員が不十分な情報（例：「Oil leak」）を入力すると、各部品の予測確率が低くなり、適切な部品選択が困難になる。そこで、入力されたテキストに基づいてLLMが推薦部品ごとの高品質テキストを生成し、保守員が意図する内容（例：「Oil leak at

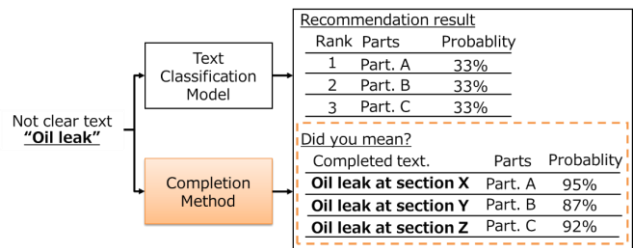


図1 入力補完の利用イメージ

section X」)に補完することで、部品の予測確率を大幅に高め、精度の高い修理対応を可能にする。

2.2 課題

関連研究として、大規模言語モデル(LLM)による入力補完手法は数多く提案されており、OpenAIのGPTシリーズ[3]やMetaのLlama3シリーズ[4]などが代表例である。これらを活用することで、文章作成支援やコード補完など多様なタスクに対応できる。LLMはWeb上の膨大なテキストを事前学習し、豊富な言語生成能力を備えているため、柔軟かつ自然な文章を生成できる点が特徴である。

関連研究では2つ課題がある。1つ目はテキスト分類モデルの予測確率が高くなるテキストに補完できないことである。LLMは大規模な言語コーパスをもとに汎用的な言語生成を行うため、特定のテキスト分類モデルの予測確率を高めるよう最適化されておらず、生成テキストが必ずしもテキスト分類モデルの予測確率を高くするとは限らない。

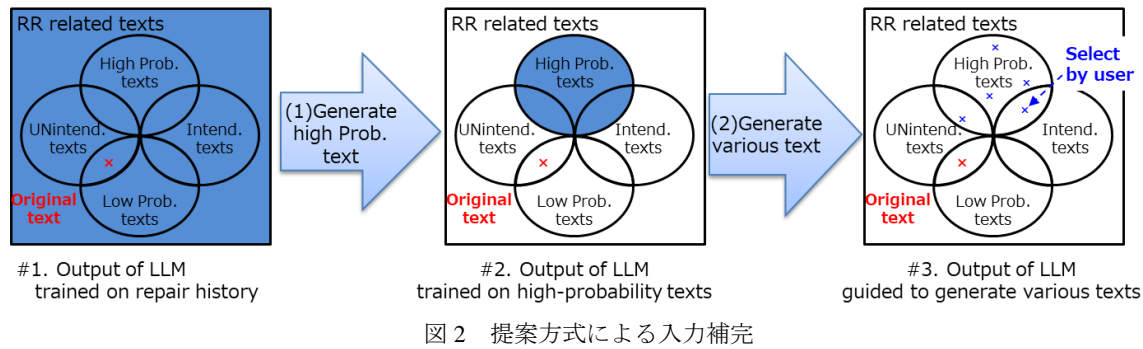
2つ目は保守員が意図するテキストの候補を生成できないことである。LLMが生成する文章は必ずしも保守員の意図を反映するとは限らないため、多様な補完テキストを生成して保守員に選択させるアプローチが考えられる。しかし、保守員が意図する可能性のあるテキスト候補を生成するには、実際に発生しうる障害の多様性を踏まえた生成が必要である。単純な単語の言い換えによる表面的な多様化だけでは、保守員が求める内容を含むテキストが生成候補に含まれる可能性は低くなる。

3. 提案方式

3.1 アプローチ

図2は、本研究の工夫を視覚的に示したもので、ベン図によって補完候補となるテキストの分布を表し、段階的に補完性能を向上させることを示している。本方式は、保守員が意図しないかつテキスト分類モデルの予測確率が低いテキスト(User-Unintended & Low prob. texts)を、保守員が意図し予測確率が高いテキスト(User-Intended & High prob. texts)へ補完することを目的とする。

[†](株)日立製作所 研究開発グループ
Hitachi, Ltd. Research & Development Group



#1 は単に修理履歴を学習させた LLM の出力範囲であり、保守員の意図や予測確率にかかわらずテキストが生成される。そこで 1 つ目の工夫として、テキスト分類モデルの予測確率が高いテキストのみを学習させることで、高確率テキストを中心に生成できるようにした (#2)。しかし、この段階で単にサンプリングするとデータの数が多き部品ばかりに偏り、多様なテキストが生成されない。そこで 2 つ目の工夫として、補完前テキストを入力した際の上位 k 件の部品ごとに LLM をガイドしてテキストを生成させる。これにより、保守員が意図する内容かつ高い予測確率をもつテキスト候補を得やすくなる。

3.2 実現方法

提案方式は 3 つの手順で実現する。

ステップ 1. データセットの準備

装置に発生した障害に関する(高品質なテキスト、低品質なテキスト、部品番号)をまとめた入力補完データセットを準備する。

ステップ 2. LLM のファインチューニング

入力補完データセットと以下の補完用プロンプトを用いて LLM をファインチューニングする。

あなたは設備保守のエキスパートです。以下の曖昧なテキストを具体的なテキストに書き換えてください。
書き換えタイプ：
{部品番号}
書き換え前：
{低品質なテキスト}
書き換え後：
{高品質なテキスト}

ステップ 3. 入力補完

低品質テキストをテキスト分類モデルに入力して上位 k 件の修理部品を取得し、(低品質テキスト、部品番号)を補完用プロンプトに当てはめてファインチューニング済み LLM で高品質テキストを生成する。この処理を k 回繰り返し、部品ごとの補完テキストを出力する。

4. 評価実験

4.1 方法

評価には、企業内で収集された非公開の修理履歴データセット約 2000 件を用いた。本データセットは、実際に装置に発生した障害に対して保守員が記載した 100 字程度のテ

キストを集めており、補完性能の妥当な検証が可能であると考えられる。このデータセットを学習したテキスト分類モデルの正解率は 47.8%であった。提案方式には Llama-3-ELYZA-JP-8B[5]を使用し、ファインチューニング前のモデルを従来方式とした。評価指標として以下の手順で得られる入力補完成功率を用いる。

1. 入力補完データセットを訓練用と評価用に 9:1 で分割し、訓練用データセットのみを LLM に学習
2. 評価用データセットの低品質なテキストに対して提案方式を適用し、5 件の補完したテキストを生成
3. 評価用データの低品質テキストに提案方式を適用し、5 件の補完テキストを生成
4. それぞれの補完テキストをテキスト分類モデルに入力し、正解部品の予測確率が 50%以上のテキストがあれば補完成功とし、評価データセット全体に対する補完成功の割合を算出

4.2 結果

表 2 に各方式の補完成功率を記載する。提案方式は従来方式と比べて高い補完成功率となった。提案方式が提示するテキストの候補の中から意図するテキストを保守員が選択することで、修理部品推薦システムの実質的な正解率を 47.8%から 65.2%に向上できる見込みを得た。

表 2. 評価実験結果

	提案方式	従来方式
入力補完成功率	33.0%	12.6%

5. おわりに

本研究では LLM を利用して低品質なテキストから高品質なテキストの候補を複数生成し、保守員に提示・選択させることで高品質なテキストを簡単に記載できる入力補完方式を提案し、従来手法と比べて入力補完成功率を向上できる見込みを得た。

今後は、非公開データに基づく評価結果の再現性を高めるため、公開可能なデータセットを用いた実験を実施する。

参考文献

- [1] 富士経済, “設備メンテナンス・ビジネス関連市場の現状と将来展望”, (2021).
- [2] 株式会社日立製作所, “部品選定支援サービス”, https://www.hitachi.co.jp/products/it/control_sys/skilled_maintenance/repair_recommend/selection.html, (2025/3/25 参照)
- [3] OpenAI, “ChatGPT”, <https://chatgpt.com/>, (2025/3/25 参照)
- [4] H. Touvron, et al. “LLaMA: Open and Efficient Foundation Language Models”, arXiv:2302.13971, (2023).
- [5] M. Hirakawa, et al. “Llama-3-ELYZA-JP-8B”, <https://huggingface.co/elyza/Llama-3-ELYZA-JP-8B>, (2024).