

Generating Semantic Matches Between Maintenance Work Orders for Diagnostic Decision Support

Syed Meesam Raza Naqvi ¹, Mohammad Ghufuran ², Safa Meraghni ³, Christophe Varnier ⁴, Jean-Marc Nicod ⁵, and Nouredine Zerhouni ⁶

^{1,3,4,5,6} *FEMTO-ST institute, Université de Bourgogne Franche-Comté / CNRS / ENSMM, Besançon, France*
syedmeesam.naqvi, safa.meraghni, christophe.varnier, jean-marc.nicod, noureddine.zerhouni@femto-st.fr

^{1,2} *Fives CortX, Lyon, France*
meesam.naqvi, mohammad.ghufuran@fivesgroup.com

ABSTRACT

Improving maintenance knowledge intelligence using text data is challenging since the maintenance information is mainly recorded as text. To unlock the knowledge from the maintenance text, a decision-making solution based on retrieving similar cases to help solve new maintenance problems is proposed. In this work, an unsupervised domain fine-tuning technique, Transformer-based Sequential Denoising Auto-Encoder (TSDAE) is used to fine-tune the BERT (Bidirectional Encoder Representations from Transformers) model on domain-specific corpora composed of the Maintenance Work Orders (MWOs). Unsupervised fine-tuning helped the BERT model to adapt MWOs text. Results indicate fine-tuned BERT model can generate semantic matches between MWOs regardless of the complex nature of maintenance text.

1. INTRODUCTION

Recently, Prognostics and Health Management (PHM) has emerged as a key technology to overcome the limitations of traditional reliability analysis. However, most of PHM research focuses on utilizing sensory signals from an engineered system to detect and diagnose faults and ignores human expertise (Atamuradov et al., 2020; Meraghni et al., 2021). Maintenance is a human knowledge-centered activity, with most activity records being textual Maintenance Work Orders (MWOs) (Bouabdallaoui, Lafhaj, Yim, Ducoulombier, & Bennadji, 2020). The MWOs contain the health history of an asset related to inspections, diagnostics, and corrective actions of the equipment reported by maintainers. Although this knowledge can be rich in contextual and system-

level practical information, few maintenance researchers attempt to exploit this human knowledge for decision making (Y. Gao, Woods, Liu, French, & Hodkiewicz, 2020). This is due to the difficulty of exploiting data contained in MWOs. The MWOs often suffer from spelling errors, domain-specific jargon, and abbreviations that prevent their immediate use in computational analyses. Consequently, retrieving similar cases from complex knowledge generated by the maintainer is challenging. Therefore, knowledge retrieval should not only retrieve similar items but also consider the context of both query and past knowledge (Brundage, Sharp, & Pavel, 2021).

As an advanced tool applied in recent years, natural language processing (NLP) allows knowledge retrieval by comparing the textual similarity between a query and stored knowledge (Akhbardeh, Desell, & Zampieri, 2020). Although many scientific works seek to adapt NLP tools to specific domains to analyze technical textual data (H. Wang, Meng, & Zhu, 2022). Technical Language Processing (TLP) is a human-in-the-loop and iterative approach to adapting NLP tools to engineering data. TLP applies engineering principles and practices to NLP to exploit the language generated by experts in their technical tasks, systems, and processes. TLP can unlock maintenance knowledge hidden in the text, providing needed insights from the asset health history while making maintenance decisions (Brundage, Sexton, Hodkiewicz, Dima, & Lukens, 2021; Liao et al., 2019).

Pre-trained language models based on Transformers, such as BERT (Bidirectional Encoder Representations from Transformers) and its variants, have achieved state-of-the-art results in several downstream NLP tasks on generic benchmark datasets (Whang et al., 2019; Z. Gao, Feng, Song, & Wu, 2019; Yu, Su, & Luo, 2019). However, BERT has been reported to underperform in specialised domains,

Syed Meesam Raza Naqvi et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 United States License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

such as biomedical (Gu et al., 2021; Cho & Lee, 2019), legal (Chalkidis, Fergadiotis, Malakasiotis, Aletras, & Androutsopoulos, 2020) and maintenance (Naqvi, Varnier, Nicod, Zerhouni, & Ghufraan, 2022). To overcome this limitation, BERT needs to be fine-tuned on domain-specific corpora. However, for most tasks and domains, labeled data is not available, and data annotation is expensive (Huang, Wei, Cui, Zhang, & Zhou, 2020). Unsupervised training approaches have been proposed to overcome this limitation: learning to embed sentences just using an unlabeled corpus for training (Han & Eisenstein, 2019).

This work focuses on developing a decision support solution for the maintenance team by exploiting the knowledge in the MWOs. Our approach is based on retrieving similar cases to help solve new maintenance problems based on past experiences to help diagnostic decision-making. To adapt the BERT model to the maintenance-specific domain, the BERT model is fine-tuned using an unsupervised domain fine-tuning technique, Transformer-based Sequential Denoising Auto-Encoder (TSDAE). The rest of this paper is organized as follows: Section 2 details the proposed methodology; Section 3 focuses on the results and discussion; The last section contains a conclusion and suggestions for future work.

2. PROPOSED METHODOLOGY

In this study, we are interested in using maintenance records to develop a decision support system based on maintenance text. The main purpose of the system is to help solve new maintenance problems based on past experiences to aid diagnostic decision-making. This section is dedicated to discussing the development of our automatic Technical Language Processing (TLP) pipeline. In this section, the dataset used for the study and various system development steps are discussed in detail, including preprocessing, domain-specific fine-tuning, and similar case retrieval policy.

2.1. Dataset

The dataset used is an open-source collection of 5485 MWOs from mining excavators over ten years (starting from 2002) (Hodkiewicz, Batsioudis, Radomiljac, & Ho, 2017). Due to the public nature of the dataset, it has been previously used in literature. It is been used to compare data-driven tagging and rules-based keyword extraction (Sexton, Hodkiewicz, Brundage, & Smoker, 2018). Sexton and Fuge utilized this dataset to extract structured information from unstructured MWOs (Sexton & Fuge, 2020). Yang et al. developed a system for the identification of degradation states of equipment using CNN (Convolutional Neural Network) based clustering (Yang, Baraldi, & Zio, 2020). In another study, authors used this dataset to explore the quality of domain-specific technical word representations using differ-

Sr.	Maintenance Work Order (MWO)
1	Bolt missing L/5 roller Bolt missing L/H No5 load roller No3 L load roller bolt broken or missing
2	SCV 1 Boily 1 Fitter 08 & 09/04 SHD24 SCV Labour 1 Fitter 1 Boily SHD 24 29/05 SCV Labour SHD24 1 Fitter 1Boily 29/09 SCV Labour 2XFitter1xBoilySHD24 13-14/04
3	L/H/Side rear estop pull cord u/s Cracked mount l/h/s R/H pump box big oil leak - R.H.S. engine - R.H. bank

Table 1. Sample Maintenance Work Orders (MWOs)

ent NLP models (Nandyala, Lukens, Rathod, & Agarwal, 2021). However, to the best of our knowledge, this dataset has not been previously used for semantic similarity. In this study, we used this dataset to find semantic similarities among MWOs using state-of-the-art NLP models. Table 1 shows some sample records from these MWOs. The cases in samples 1 and 2 show similar cases described differently by maintenance operators. Sample 3, on the other hand, shows how different operators refer to various sides of different subsystems. It can be observed in Table 1 that these samples have unconventional writing styles and are complex to process.

2.2. Preprocessing

MWOs contain domain-specific text, and this text is not easy to process using regular NLP pipelines (Brundage, Sexton, et al., 2021). MWOs contain many irregularities, such as spelling mistakes and the use of the same entity in different ways, including different spellings and acronyms. These variations combined with different writing styles of operators make it very difficult to find the semantic similarity among MWOs. Because of these reasons, the most common way of processing domain-specific text is to develop custom preprocessing pipelines. These custom pipelines are dataset-specific specialized pipelines targeted toward normalization of the text before feeding it to NLP models. The process of normalization usually involves but is not limited to spelling correction, changing acronyms to a uniform format, and conversion of short word forms to full forms throughout the dataset. These custom pipelines are normally developed with the help of domain experts using available domain knowledge. The developing process is also time-consuming and requires lot of manual labor to cover all possible scenarios. Few of the major disadvantages of these custom pipelines are the high development time, limitation to specific dataset and requirement of consistent updates as dataset grows with time. In

this study we proposed a method using state-of-the-art NLP for technical text processing which automatically handles the challenges of MWOs without the need of developing custom pipelines.

Table 1 show some samples from dataset and highlights some challenges in MWOs. For example, Sample 3 of Table 1 shows how different operators refer to spatial orientation differently. In maintenance text, these variations are shared across different terms. Also, new term variations are added with time, so the custom preprocessing pipelines need a constant upgrade. The proposed methodology handles these shortcomings of custom pipelines automatically, is quick to develop, and does not require the intervention of domain experts. Studies show that the proposed approach is also independent of a certain type of domain-specific data and can be used for a variety of technical textual records (Naqvi, Varnier, et al., 2022; Naqvi, Meraghni, et al., 2022). To observe the performance of the proposed methodology, we reduced the preprocessing to a minimum by normalizing the case to uppercase characters.

2.3. Domain specific fine-tuning

After preprocessing, to find relevant matches to the new maintenance problem, we first need a machine learning model to convert text to embeddings (numerical vectors). These embeddings can then be used to compare similarities between the textual description of the new maintenance problem (input query) and the past maintenance records. Advanced NLP models like BERT (Bidirectional Encoder Representations from Transformers) can produce semantic embeddings. BERT belongs to a class of large pre-trained models trained on a significant text source. In this study, we use BERT to covert MWO text to semantic embeddings and use it as an automatic TLP pipeline. There are many pre-trained BERT-based models. We use the “bert-base-uncased” that was proposed in the original paper on BERT (Devlin, Chang, Lee, & Toutanova, 2018).

Although pre-trained models like BERT can process text out of the box, studies have shown that fine-tuning helps when working with domain-specific data compared to regular text (Naqvi, Meraghni, et al., 2022). Also, BERT is a word embedding model and can not generate sentence or paragraph-level embedding. To address this issue, sentence BERT¹ model was proposed in 2019 (Reimers & Gurevych, 2019). The initial sentence BERT model was proposed as a supervised fine-tuning technique to adopt the BERT model for sentence/paragraph level embeddings, followed by unsupervised fine-tuning techniques (Carlsson, Gyllensten, Gogoulou, Hellqvist, & Sahlgren, 2020; T. Gao, Yao, & Chen, 2021; Janson, Gogoulou, Ylipää, Cuba Gyllensten, & Sahlgren, 2021; K. Wang, Reimers, & Gurevych, 2021).

¹<https://www.sbert.net/>

These techniques are widely tested on regular text, but there is a need to test the performance of these techniques on domain-specific text such as MWOs. As MWOs are mostly unlabeled textual descriptions of maintenance events, we use an unsupervised domain fine-tuning technique, Transformer-based Sequential Denoising Auto-Encoder (TSDAE) (K. Wang et al., 2021). TSDAE is selected as the preferred method for this study because it is the most efficient method among current unsupervised domain fine-tuning techniques with state-of-the-art performance (T. Gao et al., 2021; Janson et al., 2021; K. Wang et al., 2021). This study is mainly focused on demonstrating how state-of-the-art NLP models such like BERT can be used to leverage the underutilized source of knowledge (MWOs). A comparison of different approaches is out of the scope of this study.

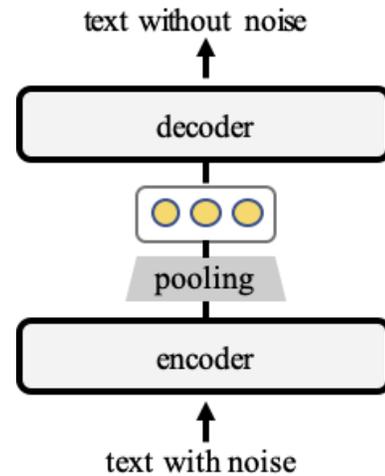


Figure 1. TSDAE Architecture (K. Wang et al., 2021)

Figure 1 shows the architecture of TSDAE which consists of an encoder (transformer model), a pooling layer, and a decoder. The first step in the fine-tuning process is the development of a noisy dataset. In this step noise is added to sentences of the original dataset. In this step noise could be for example deleting certain words from the original sentence. After adding noise to all the sentences in the dataset these sentences are fed to a transformer model (BERT in our case) which acts as an encoder and outputs token embeddings. These token embeddings are then converted to single fixed-size sentence embedding using a pooling layer. The task of the decoder layer is to generate the original sentence using the noisy input. During fine-tuning process weights of the encoder are optimized based on the response of the decoder. After fine-tuning only encoder is used to generate sentence or paragraph-level embeddings. TSDAE takes a pre-trained transformer-based model and fine-tunes the models using domain-specific data without any labels. The output of the process is a domain fine-tuned sentence BERT model trained

in an unsupervised manner that can generate fixed-length sentence embeddings. As TSDAE focuses on fixed-size sentence embeddings instead of token embeddings during fine-tuning, the resulting sentence embedding model gives better sentence and paragraph level embeddings than the pre-trained version of the input model. The size of the final fixed-size sentence embedding is the same as BERT embedding size (768 for BERT base model). TSDAE can be considered as a modified encoder-decoder transformer in which the key and value of cross-attention are focused on fixed-size sentence embeddings. Equation 1 and 2 show the formulation of modification in cross-attention.

$$H^{(k)} = \text{Attention}\left(H^{(k-1)}, [s^T], [s^T]\right) \quad (1)$$

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d}}\right)V \quad (2)$$

In equation 1, $H^{(k)} \in \mathbb{R}^{t \times d}$ is hidden states of decoder within t decoding steps at k -th layer, $[s^T] \in \mathbb{R}^{1 \times d}$ is a single row matrix of sentence embedding vector and d is the size of the sentence embedding vector. Q, K and V in equation 2 are cross-attention query, key and value respectively (K. Wang et al., 2021). In TSDAE paper authors explored various methods of adding noise to the sentence but deletion with a deletion ratio of 0.6 resulted in the best performance.

Figure 2 shows the pipelines of unsupervised domain fine-tuning of the BERT model using TSDAE to generate the sentence embedding model. The first step is the preprocessing, where we only normalized the words present in input MWOs, as explained in the preprocessing section. After preprocessing, the TSDAE fine-tunes pre-trained BERT model using domain-specific data, and the output of the process is a domain-specific sentence BERT model.

2.4. Retrieval

After fine-tuning the BERT model on MWOs, we can now use this model to retrieve similar MWOs to the input query text. Figure 3 shows steps involved in retrieving top k similar maintenance work orders, where k represents number of similar cases retrieved. As presented in Figure 3, in the case of a new maintenance problem, the description of the problem is first converted to semantic embedding. Afterward, we extract the embeddings of past MWOs and compare them with the embedding of the input query using cosine similarity. Equation 3 shows this cosine similarity which is the measure of similarity between two non-zero vectors \mathbf{A} and \mathbf{B} .

$$\text{cosine similarity} = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\| \times \|\mathbf{B}\|} \quad (3)$$

Finally, retrieved similar cases can be used as reference past

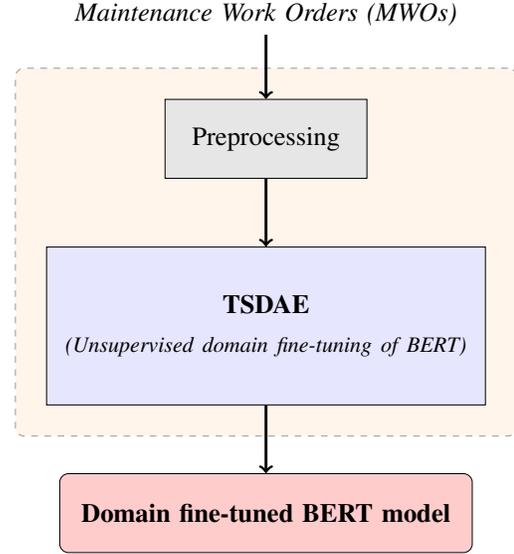


Figure 2. Unsupervised domain fine-tuning of BERT model using TSDAE to generate sentence embedding model

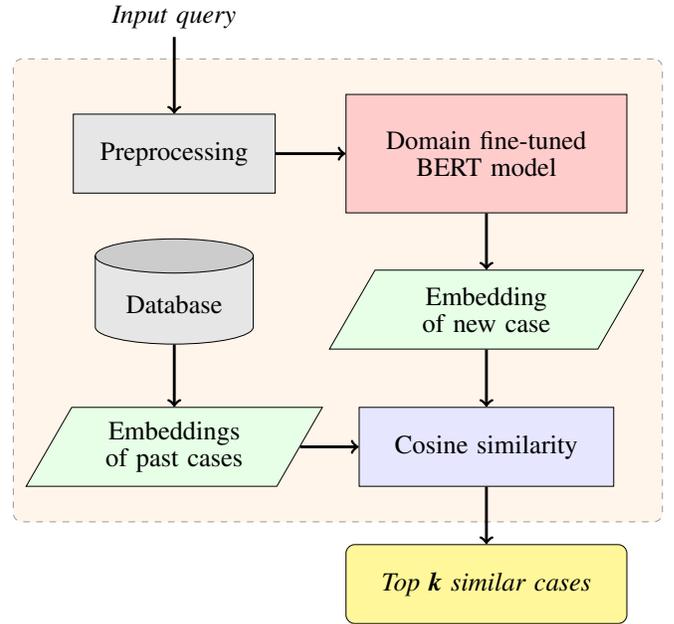


Figure 3. Retrieval of top k similar Maintenance Work Orders (MWOs) to the input query

cases to help solve new maintenance problems described in the input query. To make sure that retrieved cases are relevant, we only retrieved cases with 80% or more similarity and set the value of k to at most 5 cases. We also ensured no duplication in retrieved similar cases. Afterward, we analyzed the cases for some interesting patterns, and some of the cases from the analysis are presented in the results section.

3. RESULTS AND DISCUSSION

In this section, the results of the study are discussed. For each case in the database, we retrieved similar cases given the retrieval criteria, to test the performance of developed domain fine-tuned model. Then we categorized retrieved results into one or more categories based on the nature of the match. There are two main categories: exact match and semantic match. The semantic match is further categorized into three subcategories: (1) entity linkage, (2) similar meaning, and (3) spelling coverage. These subcategories of semantic match correspond to the challenges in MWOs. Samples in the dataset are raw MWOs but we categorized the results to demonstrate the performance of the proposed methodology. The reason behind evaluating the results based on these categories is the correspondence of these categories to the challenges of processing MWOs (such as spelling mistakes, use of similar entities in different ways, and dataset instances that have similar meanings). Each category with its subcategories is described below:

- **Exact match:** We categorize a retrieved case as an exact match when it is the same as the input query or contains part of the input query;
- **Semantic match:** We categorize a retrieved case as a semantic match if it has some semantic patterns such as similar meaning, entity linkage, or spelling mistake coverage:
 - **Entity linkage:** We categorize a retrieved case as entity linkage when the model can identify an entity from an input query used differently in the retrieved case. For example, in the case of the location of a system/subsystem, the right hand can be represented by “RH” in the input query. However, it may be represented differently in retrieved cases such as “R/H”, “R/H/Side”, or “RHS”;
 - **Similar meaning:** We categorize a retrieved case as having similar meaning when the model can retrieve a case that is similar in meaning but not an exact match to the input query;
 - **Spelling coverage:** We categorize a retrieved case in the spelling coverage category if the model can identify words from the input query in a retrieved similar case regardless of spelling mistakes.

Table 2 shows the results of nine input queries with retrieved similar cases with assigned categories. Interesting patterns in retrieved cases for each query are individually discussed below:

- **Query 1:** For query 1, all the retrieved cases fall into the similar meaning category. It is interesting to see how the model can match patterns like “2 OFF” with “2”, “2 LOST”, “BROKEN X 2”, and “BUCKET TEETH 3 & 4” showing efficient semantics matching;

- **Query 2:** We only got four similar cases based on retrieval criteria for query 2. Retrieved cases fall into different categories. It is interesting to see that for “AIRCONS (Air Conditioner) NOT GETTING COLD” model can identify “AIRCON COMPRESSORS NOT WORKING” as a related case. Similarly, the model can identify the complete form (AIRCONDITIONER) from the short form (AIRCONS) even with spelling mistakes (AIR and CONDITIONER are merged), showing entity linkage. Also, the model can identify “SHD24” and “SHD0024” as the same entities;
- **Query 3:** In query 3, we have entity linkage between “GREASE LINES” and “LUBE LINES”. We also have some cases that fall in the similar meaning category and exact match even with the spelling mistake (LIN instead of LINES);
- **Query 4:** For query 4, we have cases with exact matches and similar meanings. We also have entity linkage (between SDH 24 and SHD24). The model also identified a similar case with the spelling mistake (FITT instead of FITTER);
- **Query 5:** Query 5 indicates how different operators describe the short form of various system/subsystem sides. For example, RH is used to indicate the right-hand side in the query, and the model can identify “R/H”, “L/H”, and “RHS” are similar entities in retrieved similar cases;
- **Query 6:** For query 6, we only got three similar cases. Results of query 6 are interesting as the model can identify a complex shorthand pattern “L/H NO5 LOAD ROLLER”, which means load roller number 5 on the left-hand side. In similar cases, the model can identify patterns like “L/5 (left-hand number 5)”, “NO3 L (left-hand number 3)”, and “R # 1 (roller number 1)”;
- **Query 7:** In query 7, from the statement “LUBE PUMP HAS FAILED” the model can identify cases with similar meanings such as “LUBE SYSTEM FAILURE” and “LUBE SYSTEM FAILURE”;
- **Query 8:** For query 8, we have only two similar cases. Both are an exact match, but the second one is an exact match with spelling coverage (OERHEATING instead of OVERHEATING);
- **Query 9:** Finally, in query 9, we only got one similar case based on retrieval criteria. From the input query “CRACKS UNDER HEEL OF BUCKET” model can identify a similar case with “CRACKING ON BOTTOM OF BUCKET” that falls into a similar meaning category and entity linkage (HEEL OF BUCKET matched with BOTTOM OF BUCKET).

Some common interesting patterns in query and similar cases text are highlighted using bold letters in Table 2. Results presented in Table 2, indicate that given an input query developed domain fine-tuned model without using custom pipelines can

Table 2. Categorizing similar case based on similarity patterns

Sr.	Input Query	Similar Cases	Semantic match			Exact match	Similarity (%)
			Entity linkage	Similar meaning	Spelling coverage		
1	REPLACE BUCKET TEETH 2 OFF	REPLACE 2 BUCKET TEETH		■			93.7
		REPLACE 2 BUCKET TEETH AND KEEPERS		■			86.1
		REPLACE 2 LOST TEETH		■			83.3
		BUCKET TEETH BROKEN X 2		■			80.8
		REMOVE BUCKET TEETH 3 & 4		■			80.5
2	AIRCONS NOT GETTING COLD SHD24	AIRCONS NOT GETTING COLD SHD24				■	100
		AIRCON COMPRESSORS NOT WORKING SHD24	■	■			87.4
		AIRCON NOT WORKING SHD0024	■	■			83.8
		AIRCONDITIONER NOT GETTING COLD	■	■	■		83
3	REPLACE GREASE LINES TO BUCKET PINS	REPLACE GREASE LINES TO BUCKET PINS				■	100
		REPLACE LUBE LINES ON BUCKET PINS	■	■			90
		REPAIR GREASE LINES ON BUCKET PIN		■			88.9
		REPLACE GREASE LINES ON BUCKET		■			86.1
		REPLACE GREASE LIN TO BUCKET PIN			■	■	85.4
4	SCV LABOUR 1 FITTER 1 BOILY SHD 24 29/05	SCV LABOUR 1 FITTER 1 BOILY SHD 24 16/06				■	99.1
		SCV LABOUR 1 FITTER 1 BOILY EXD 36 06/05		■			97.3
		SCV LABOUR SHD24 I FITTER 1 BOILY 27/10	■	■			96.4
		SCV LABOUR SHD24 I FITT 1 BOILY 03/12/09	■	■	■		95.5
		SCV 1 BOILY 1 FITTER 08 & 09/04 SHD24		■			90.9
5	REPLACE RH BOOM CYLINDER	REPLACE RH BOOM CYLINDER				■	100
		REPLACE R/H BOOM CYLINDER	■			■	90.1
		REPLACE L/H BOOM CYLINDER	■			■	88.6
		RHS BOOM CYLINDER LEAKING	■	■			86.7
		SHD24 REPLACE R/H BOOM CYLINDER	■	■			82.9
6	BOLT MISSING L/H NO5 LOAD ROLLER	BOLT MISSING L/5 ROLLER	■	■			88.8
		NO3 L LOAD ROLLER BOLT BROKEN OR MISSING	■	■			88.5
		R # 1 LOAD ROLLER MISSING 1 BOLT	■	■			88.5
7	UNDERCARRIAGE LUBE PUMP HAS FAILED	UNDERCARRIAGE LUBE SYSTEM FAILURE		■			85.2
		UNDERCARRIAGE LUBE NOT WORKING		■			80
8	R/H ENGINE OVERHEATING SEE TEXT RAN TEM	R/H ENGINE OVERHEATING				■	81.8
		ENGINE OERHEATING SEE TEXT		■	■		80.5
9	CRACKS UNDER HEEL OF BUCKET	REPAIR CRACKING ON BOTTOM OF BUCKET	■	■			85.1

efficiently identify similar cases regardless of variation in the entity names, way of description, and spelling mistakes.

4. CONCLUSION AND FUTURE WORK

In this study, an automatic TLP pipeline is developed using the BERT model. Results indicate that with proper fine-tuning, state-of-the-art models such as BERT can efficiently process domain-specific text such as MWOs. The developed system can identify complex semantic patterns among different MWOs. Retrieved similar cases through the system against various queries indicated for most cases, the fine-tuned model was able to identify similar cases with complex patterns. Retrieved similar cases through the model show diversity in coverage for various categories, including exact match, entity linkage, spelling mistake handling, and similar meaning cases. One of the limitations of this study is the need for periodic model fine-tuning as new terms and cases are introduced, but that is also true for other techniques such as manual TLP pipelines to cover various scenarios. However, given the automatic nature of our proposed process, the effort required to update a model is considerably lesser than manual pipelines.

Some of the future work planned for this case study includes: (i) identifying new ways to measure the quality of semantic similarity in retrieved cases; (ii) analyzing the performance of the developed framework on other unsupervised fine-tuning approaches in the literature; (iii) formalize the proposed system as a service in industrial manufacturing setup.

5. ACKNOWLEDGEMENT

This work has been supported by the EIPHI graduate school (contract “ANR-17-EURE-0002”).

REFERENCES

- Akhbardeh, F., Desell, T., & Zampieri, M. (2020). Nlp tools for predictive maintenance records in maintnet. In *Proceedings of the 1st conference of the asia-pacific chapter of the association for computational linguistics and the 10th international joint conference on natural language processing: System demonstrations* (pp. 26–32).
- Atamuradov, V., Medjaher, K., Camci, F., Zerhouni, N., Derzin, P., & Lamoureux, B. (2020). Machine health indicator construction framework for failure diagnostics and prognostics. *Journal of Signal Processing Systems*, 92(6), 591–609.
- Bouabdallaoui, Y., Lafhaj, Z., Yim, P., Ducoulombier, L., & Bennadji, B. (2020). Natural language processing model for managing maintenance requests in buildings. *Buildings*, 10(9), 160.
- Brundage, M. P., Sexton, T., Hodkiewicz, M., Dima, A., & Lukens, S. (2021). Technical language processing: Unlocking maintenance knowledge. *Manufacturing Letters*, 27, 42–46.
- Brundage, M. P., Sharp, M., & Pavel, R. (2021). Qualifying evaluations from human operators: Integrating sensor data with natural language logs. In *Phm society european conference* (Vol. 6, pp. 9–9).
- Carlsson, F., Gyllensten, A. C., Gogoulou, E., Hellqvist, E. Y., & Sahlgren, M. (2020). Semantic re-tuning with contrastive tension. In *International conference on learning representations*.
- Chalkidis, I., Fergadiotis, M., Malakasiotis, P., Aletras, N., & Androutsopoulos, I. (2020). Legal-bert: The muppets straight out of law school. *arXiv preprint arXiv:2010.02559*.
- Cho, H., & Lee, H. (2019). Biomedical named entity recognition using deep neural networks with contextual information. *BMC bioinformatics*, 20(1), 1–11.
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.
- Gao, T., Yao, X., & Chen, D. (2021). Simcse: Simple contrastive learning of sentence embeddings. *arXiv preprint arXiv:2104.08821*.
- Gao, Y., Woods, C., Liu, W., French, T., & Hodkiewicz, M. (2020). Pipeline for machine reading of unstructured maintenance work order records. In *Proceedings of the 30th european safety and reliability conference and 15th probabilistic safety assessment and management conference (esrel)*.
- Gao, Z., Feng, A., Song, X., & Wu, X. (2019). Target-dependent sentiment classification with bert. *Ieee Access*, 7, 154290–154299.
- Gu, Y., Tinn, R., Cheng, H., Lucas, M., Usuyama, N., Liu, X., ... Poon, H. (2021). Domain-specific language model pretraining for biomedical natural language processing. *ACM Transactions on Computing for Healthcare (HEALTH)*, 3(1), 1–23.
- Han, X., & Eisenstein, J. (2019). Unsupervised domain adaptation of contextualized embeddings for sequence labeling. *arXiv preprint arXiv:1904.02817*.
- Hodkiewicz, M. R., Batsioudis, Z., Radomiljac, T., & Ho, M. T. (2017). Why autonomous assets are good for reliability—the impact of ‘operator-related component’ failures on heavy mobile equipment reliability. In *Annual conference of the phm society* (Vol. 9).
- Huang, S., Wei, F., Cui, L., Zhang, X., & Zhou, M. (2020). Unsupervised fine-tuning for text clustering. In *Proceedings of the 28th international conference on computational linguistics* (pp. 5530–5534).
- Janson, S., Gogoulou, E., Ylipää, E., Cuba Gyllensten, A., & Sahlgren, M. (2021). Semantic re-tuning with contrastive tension. In *International conference on learning representations, 2021* (p. 1-21).

- Liao, W., Yin, Y., Ye, G., Zuo, D., Zhang, Q., & Ma, Y. (2019). Automobile maintenance case matching method based on the continuous space language model. In *2019 international conference on economic management and model engineering (icemme)* (pp. 476–481).
- Meraghni, S., Terrissa, L. S., Yue, M., Ma, J., Jemei, S., & Zerhouni, N. (2021). A data-driven digital-twin prognostics method for proton exchange membrane fuel cell remaining useful life prediction. *International Journal of Hydrogen Energy*, *46*(2), 2555–2564.
- Nandyala, A. V., Lukens, S., Rathod, S., & Agarwal, P. (2021). Evaluating word representations in a technical language processing pipeline. In *Phm society european conference* (Vol. 6, pp. 17–17).
- Naqvi, S. M. R., Meraghni, S., Varnier, C., Nicod, J.-M., Zerhouni, N., & Ghufran, M. (2022). Cbr-based decision support system for maintenance text using nlp for an aviation case study. In *2022 prognostics and health management conference (phm-london)* (pp. 344–349).
- Naqvi, S. M. R., Varnier, C., Nicod, J.-M., Zerhouni, N., & Ghufran, M. (2022). Leveraging free-form text in maintenance logs through bert transfer learning. In *International conference on deep learning, artificial intelligence and robotics* (pp. 63–75).
- Reimers, N., & Gurevych, I. (2019). Sentence-bert: Sentence embeddings using siamese bert-networks. *arXiv preprint arXiv:1908.10084*.
- Sexton, T., & Fuge, M. (2020). Organizing tagged knowledge: Similarity measures and semantic fluency in structure mining. *Journal of Mechanical Design*, *142*(3).
- Sexton, T., Hodkiewicz, M., Brundage, M. P., & Smoker, T. (2018). Benchmarking for keyword extraction methodologies in maintenance work orders. In *Proceedings of the annual conference of the phm society* (Vol. 10).
- Wang, H., Meng, X., & Zhu, X. (2022). Improving knowledge capture and retrieval in the bim environment: Combining case-based reasoning and natural language processing. *Automation in Construction*, *139*, 104317.
- Wang, K., Reimers, N., & Gurevych, I. (2021). Tsdae: Using transformer-based sequential denoising auto-encoder for unsupervised sentence embedding learning. *arXiv preprint arXiv:2104.06979*.
- Whang, T., Lee, D., Lee, C., Yang, K., Oh, D., & Lim, H. (2019). An effective domain adaptive post-training method for bert in response selection. *arXiv preprint arXiv:1908.04812*.
- Yang, Z., Baraldi, P., & Zio, E. (2020). A novel method for maintenance record clustering and its application to a case study of maintenance optimization. *Reliability Engineering & System Safety*, *203*, 107103.
- Yu, S., Su, J., & Luo, D. (2019). Improving bert-based text classification with auxiliary sentence and domain knowledge. *IEEE Access*, *7*, 176600–176612.

BIOGRAPHIES



Syed Meesam Raza Naqvi is currently a Ph.D. student at the University of Franche-Comté in Besançon and a research and development engineer in data science at Fives CortX, Lyon, France. He received his B.S. degree in Telecommunication and Networking from COMSATS University Islamabad, Pakistan in 2012, and his M.S. degree in Computer Science from the National University of Science and Technology (NUST), Islamabad, Pakistan in 2019. His main research interests are Prognostic and Health Management (PHM), machine learning, Natural Language Processing (NLP), maintenance decision-support, and speech recognition.



Mohammad Ghufran obtained his Ph.D. in 2018 at Paris-Saclay University, France. He is currently responsible for Research and Development of Predictive Maintenance methods at Fives CortX for various industries. His research interests are Prognostic and Health Management (PHM) and Natural Language Processing (NLP).



Safa Meraghni is currently a Post-Doc at University of Bourgogne Franche-Comté. She obtained her Ph.D. degree in 2021 at University of Biskra, Algeria. Safa Meraghni is working on the application of artificial intelligent in industry and medical fields. She is recently focus on extracting information from narrative texts using the latest technologies of Natural Language Processing.



Christophe Varnier is an Associate Professor at ENSMM, France. He obtained his Ph.D. degree in 1996 at the University of Franche-Comté in Besançon, France. He is teaching computer science at Ecole Nationale Supérieure de Mécanique et des Microtechniques (ENSMM) since 1996. He is a Researcher at the Automatic Control and Micro-Mechatronic Systems Department of the FEMTO-ST Institute. His research interests include PHM, operation research, scheduling and optimization.



Jean-Marc Nicod obtained his Ph.D. degree in 1997, from the Ecole Normale Supérieure of Lyon (LIP at ENS) and his HDR from the Université de Franche-Comté (LIFC) in 2006. He is currently a full Professor in computer science at the Automatic Control and Micro-Mechatronic Systems Department (AS2M) of Institut FEMTO-ST in the National Engineering Institute in Mechanics and Microtechnologies (ENSMM) in Besançon, France. His research interests include operational research, optimization, algorithms, scheduling, decision making and parallel and distributed computing.



Nouredine Zerhouni received his engineer degree from National Engineers and Technicians School of Algiers (ENITA) in 1985. He received his Ph.D. Degree in Automatic Control from the Grenoble National Polytechnic Institute in 1991. In September 1991, he

joined the National Engineering School of Belfort (ENIB) as Associate Professor. Since September 1999, Nouredine Zerhouni is Professor at Ecole Nationale Supérieure de Mécanique et des Microtechniques (ENSMM) in Besançon. His main research activities are concerned with intelligent maintenance systems and e-maintenance.