MEDINFO 2023 — The Future Is Accessible
J. Bichel-Findlay et al. (Eds.)

doi:10.3233/SHTI231055

Extracting Symptoms of Agitation in Dementia from Free-Text Nursing Notes Using Advanced Natural Language Processing

Dinithi VITHANAGE[a], Yunshu ZHU[a], Zhenyu ZHANG[a], Chao DENG[b], Mengyang YIN[c] and Ping YU[a,1]
[a]*Center for Digital Transformation, School of Computing and Information Technology, University of Wollongong, Wollongong, Australia.*
[b]*School of Medical, University of Wollongong, Wollongong, Australia.*
[c]*Opal Healthcare, Sydney, Australia.*
ORCiD ID: Dinithi Vithanage https://orcid.org/0000-0001-5851-7158, Yunshu Zhu https://orcid.org/0000-0003-2786-0775, Zhenyu Zhang https://orcid.org/0000-0003-1853-4978, Chao Deng https://orcid.org/0000-0003-1147-5741, Mengyang Yin https://orcid.org/0000-0002-0212-4598, Ping Yu https://orcid.org/0000-0002-7910-9396

Abstract. Nursing staff record observations about older people under their care in free-text nursing notes. These notes contain older people's care needs, disease symptoms, frequency of symptom occurrence, nursing actions, etc. Therefore, it is vital to develop a technique to uncover important data from these notes. This study developed and evaluated a deep learning and transfer learning-based named entity recognition (NER) model for extracting symptoms of agitation in dementia from the nursing notes. We employed a Clinical BioBERT model for word embedding. Then we applied bidirectional long-short-term memory (BiLSTM) and conditional random field (CRF) models for NER on nursing notes from Australian residential aged care facilities. The proposed NER model achieves satisfactory performance in extracting symptoms of agitation in dementia with a 75% F1 score and 78% accuracy. We will further develop machine learning models to recommend the optimal nursing actions to manage agitation.

Keywords. Named entity recognition, natural language processing, deep learning, transfer learning, nursing notes, symptoms, agitation in dementia

1. Introduction

Accurate and complete observation of the health conditions of older people is the prerequisite for delivering high-quality nursing care. These observations are often recorded in the free-text nursing notes [1]. Nursing notes are a type of electronic health record that contains information about older people's disease status, symptoms, frequency of symptom occurrence, medication, and outcome of the treatments.

[1] Corresponding Author: Ping Yu, email: ping@uow.edu.au, Tel: 61 479086978.

Health data analytics has long been challenged to automatically process free-text data, e.g., observations captured in the nursing notes. The challenge can be addressed by advancement in natural language processing (NLP), in particular, the application of deep learning and transfer learning to NLP tasks. Named entity recognition is a task in NLP that automatically recognizes and labels entities [2], e.g., agitation in dementia. It is one of the important applications of NLP for classifying entities with specific meanings in a given text and can be applied to recognize the key observations in the nursing notes. Therefore, we aim to develop a NER model for extracting symptoms of agitation in dementia from the nursing notes. We believe this knowledge is important for developing person-centered treatment and care for older people with agitation in dementia in residential aged care facilities (RACFs).

Older people with dementia often manifest agitated behaviors with a variety of non-purposeful physical and verbal behaviors: resisting, wandering, repetitive questioning, cursing, screaming, hitting, shouting, spitting, and destroying property [3]. Due to the prevalence of the symptom and its significant physical and psychological strain on older people with the condition and their caregivers, family and friends, managing agitation symptoms in dementia has been an ongoing challenge that all caregivers need to face and manage in the Australian RACFs [3]. Identifying symptoms of agitation and effective nursing actions to manage the symptoms is important.

Previous work on NER for health data analytics can be classified into three classes: (1) dictionary and rule-based approach, (2) feature-based approach, and (3) neural network-based approach. The traditional NLP methods for NER include dictionary-based and rule-based approaches [3,4]. However, developing a complete vocabulary with target-specific rules requires a lot of expert human effort. Further, it is difficult to compile an extensive list of terms or rules because of the constantly evolving biomedical knowledge. The advanced NLP methods for NER include word embedding models as a feature-based approach [2,5-7]. These models used simple linear layers with softmax functions and CRF models as classification methods for NER. Further, the word embedding models used Long Short-Term Memory for handling NER as a neural network approach has shown improved performance [5-7]. Although the existing studies offer useful features for NER, they primarily rely on pre-defined features that aim to capture the unique surface properties of entity types. To date, there is little research that conducts NER in free-text nursing notes.

The contextual word embedding models can generate specific word embeddings for a word to capture its placement in a phrase and context; it is context-dependent [8]. The recent advancement in contextual word embedding is transfer learning, which applies a pre-trained model with embeddings learned from training on large datasets to a small data set to address certain NLP tasks [8]. To date, there is no specific research for transfer learning and deep learning models on NER tasks for nursing notes on agitation in dementia. The current research gap has triggered us to develop the NER model to extract symptoms of agitation in dementia from nursing notes in RACFs.

2. Methods

2.1. Ethics Approval

The University of Wollongong's Human Research Ethics Committee gave its approval to the current study (Ethics Number 2019/159).

2.2. NER Model Architecture

The proposed NER model has three layers; the first layer is the word embedding layer. The output of the word embedding layer is fed to the second layer, the relationship capture layer. The output of the second layer is fed to the third layer, the label prediction layer.

2.2.1. Word Embedding Layer

The task of this layer is word embedding, which maps different words to high-dimensional feature vectors. Clinical BioBERT is developed by pre-training the BioBERT model on clinical notes and fine-tuned for biomedical inference tasks and NER; thus, the Clinical BioBERT can learn good representations for the nursing notes to generate context-based feature vectors [9]. Therefore, we selected the Clinical BioBERT model for this task.

2.2.2. Relationship Capture Layer

The task of the second layer is to capture relationships between the feature vectors generated from the word embedding layer. The BiLSTM model processes each sequence of feature vectors from two directions, both left-to-right and right-to-left [2,5-7]. Therefore, it can learn long-term bidirectional dependencies among feature vectors and thus is the appropriate model for this task.

2.2.3. Label Prediction Layer

The task of the third layer is to capture the best label, i.e., the optimal predicted label, by learning the order dependency and the transition probability between labels from the output of the relationship capture layer [2,5-7]. The CRF model computes the joint probability distribution of the whole label sequence when an observation sequence intended for the label is available rather than making independent labelling decisions; thus, CRF is the appropriate model for this task [2,5-7].

2.3 Dataset

The dataset, nursing notes, were collected from 40 RACFs in New South Wales, Australia, in 2019 and 2020. The dataset contains 1,000 labelled nursing notes, which domain experts annotated. We adopted the Begin-Inside-Outside (BIO) tagging system to label the entities. A token is labelled as B-label if it is the first of a named entity, I-label if it is inside a named entity but not the first, and O-label otherwise.

2.4 NER Model Development

We randomly divided the datasets into training, validation, and test sets of 80%, 10%, and 10%, respectively. We trained the model using the PyTorch framework implemented on an NVIDIA V100 GPU with 32 GB of memory. We used a set of reasonable hyper-parameters to train the model, which included 32 batch size, 2e-5 learning rate with AdamW optimizer, 20 epochs, 128 hidden units of bi-LSTM, and 512 maximum sequence length, which were selected based on the existing works [2,5,6]. We used early stopping to avoid overfitting. We trained the model for each set of hyper-parameters and chose the model that performed the best on the validation set. We assessed their performance on the test set. As a baseline for our proposed NER model, we use the NER model only embedded with Clinical BioBERT as the comparison model. It classifies entities with softmax function on the same test set.

3. Results

We compared the performance of our three-layer model with the base NER model by the mean score in recognition of the symptoms of agitation in dementia as entities. Our three-layer model performed better, with an F1 score of 75% and an accuracy of 78%, compared with the base model with an F1 score of 68% and an accuracy of 70%.

4. Discussion

The provision of tailored, person-centred nursing care has been put at risk due to a lack of clear understanding of the various manifestations of agitated behaviour of older people with dementia in RACFs [3]. Therefore, it is crucial to develop an automated technique to uncover this information from the nursing notes. Currently, there is little study about the extraction of agitated behaviour from nursing notes. The previous study extracted agitated behaviour manually from the nursing notes, which is both labour-intensive and subject to self-reporting bias [10]. Another similar study used a rule-based method to automatically extract agitated behaviour from the nursing notes, which can give wrong suggestions once a word expressing the symptom is not included in the training data set, an out-of-vocabulary problem [3].

To date, there is no specific study on agitation in nursing notes to address the limitations of the existing studies. Therefore, this study provides an advanced NLP-based model solution for nursing notes to extract useful clinical information about agitation symptoms. The proposed NER model on agitation symptoms made several contributions by applying transfer learning, and deep learning in advanced NLP: (1) contributing to improving performances of NER on symptoms of agitation in dementia by leveraging the knowledge of nursing notes learned in advance; (2) contributing to accurately capture the semantics for NER even though the words are not in the training data; (3) contributing to overcome the issues of label data scarcity and need of more hardware resources to train the machine learning models to extract symptoms of agitation; (4) contributing to deal with the complex characteristics of word use under different linguistic contexts in the nursing notes, and (5) contributing to extract the agitation prevalence information from the nursing notes. The information is useful for the nursing staff to implement person-centered nursing interventions to manage and prevent agitation and improve the quality of care for older people with dementia in RACFs.

The proposed three-layer NER model outperformed the model that is only based on the transfer learning model, Clinical BioBERT. Our proposed model takes advantage of the BiSLTM and CRF models in deep learning on top of the Clinical BioBERT. Its better performance than the pure Clinical BioBERT model may be attributed to the deep learning models that can capture longer dependencies for sequence labelling in the nursing notes. Although the proposed NER model achieves satisfactory performance in extracting symptoms of agitation, this study has two limitations. The first is the number of labelled data that we used to train the model is only 1000 nursing notes. Even though the model could train on much less labelled data due to transfer learning, there is a possibility to improve the model's performance by training more labelled nursing notes. Secondly, we only conducted the NER model to extract symptoms of agitation in dementia. In the future, we plan to extract nursing actions to manage these symptoms.

5. Conclusions

We reported a NER model to extract symptoms of agitation in dementia from free-text nursing notes in RACFs. The combination of deep learning (BiLSTM and CRF) and transfer learning (Clinical BioBERT) has allowed us to improve the performance of the NER task. This study showcases the usefulness of NLP underpinned by deep learning and transfer learning for extracting key variables from free-text nursing notes. In the future, we plan to improve the model accuracy by training more labelled data. Further, we plan to extend the NER model to extract the causal factors and nursing actions to prevent agitation in dementia. We will combine the variables and values extracted from the free-text nursing notes with the structured data to conduct further machine learning to uncover the patterns or trends that will help nursing care decision-making.

References

[1] Wang N, Hailey D, Yu P. Quality of nursing documentation and approaches to its evaluation: a mixed-method systematic review. J Adv Nurs. 2011 Sep;67(9):1858-75, doi: 10.1111/j.1365-2648.2011.05634.x.

[2] Yu X, Hu W, Lu S, Sun X, Yuan Z. BioBERT based named entity recognition in electronic medical record. In2019 10th international conference on information technology in medicine and education (ITME) 2019 Aug 23 (pp. 49-52). IEEE, doi: 10.1109/ITME.2019.00022.

[3] Zhu Y, Song T, Zhang Z, Deng C, Alkhalaf M, Li W, Yin M, Chang HCR, Yu P. Agitation Prevalence in People With Dementia in Australian Residential Aged Care Facilities: Findings From Machine Learning of Electronic Health Records. J Gerontol Nurs. 2022 Apr;48(4):57-64. doi: 10.3928/00989134-20220309-01.

[4] Cook HV, Pafilis E, Jensen LJ. A dictionary-and rule-based system for identification of bacteria and habitats in text. InProceedings of the 4th BioNLP shared task workshop 2016 Aug (pp. 50-55).

[5] Sun C, Yang Z, Wang L, Zhang Y, Lin H, Wang J. Biomedical named entity recognition using BERT in the machine reading comprehension framework. J Biomed Inform. 2021 Jun;118:103799, doi: 10.1016/j.jbi.2021.103799.

[6] Liu S, He T, Dai J. A Survey of CRF Algorithm Based Knowledge Extraction of Elementary Mathematics in Chinese. Mob Netw Appl. 2021;26(5):1891-903, doi: 10.1007/s11036-020-01725-x.

[7] Giorgi JM, Bader GD. Transfer learning for biomedical named entity recognition with neural networks. BIOINFORMATICS. 2018;34(23):4087-94, doi: 10.1093/bioinformatics/bty449.

[8] Krishnan A, Das M, Bendre M, Yang H, Sundaram H, editors. Transfer Learning via Contextual Invariants for One-to-Many Cross-Domain Recommendation. Annual ACM Conference on Research and Development in Information Retrieval; 2020; Ithaca: ACM.

[9] Alsentzer E, Murphy JR, Boag W, Weng WH, Jin D, Naumann T, McDermott M. Publicly Available Clinical BERT Embeddings. 2019, doi: 10.48550/arxiv.1904.03323.

[10] Kverno KS, Rabins PV, Blass DM, Hicks KL, Black BS. Prevalence and treatment of neuropsychiatric symptoms in advanced dementia. J Gerontol Nurs. 2008;34(12):8-17, doi: 10.3928/00989134-20081201-03.