Analysis of Nursing Records Using Text Vectorization Techniques

Takahiro Kubo
Facility of Data Science, Facility of Nursing Musashino University
Tokyo, Japan
ta_kubo@musashino-u.ac.jp

Virach Sornlertlamvanich
Facility of Data Science
Musashino University
Tokyo, Japan
httpss://orcid.org/0000-0002-6918-8713

Thatsanee Charoenporn
Facility of Data Science
Musashino University
Tokyo, Japan
https://orcid.org/0000-0002-9577-9082

***Abstract*—In recent years, deep learning technology has made a significant contribution to solving various challenges in the healthcare field. For example, improvements in disease diagnosis accuracy and process reliability through image recognition and the analysis of medical records in the combination with natural language processing (NLP) techniques can be cited. This study proposes a novel method for automating the nursing diagnosis process using NLP techniques, aiming to address one of the causes of nurses' excessive workloads, especially in the case of nursing record documentation. Nursing records are written in Japanese, and the lack of spaces between words and the inclusion of patients' personal information are major challenges. This paper presents a perspective on constructing a model to assist nurses in the nursing diagnosis process, utilizing BERT techniques to realize a system that reduces nurses' workload and improves the quality of patient care. The approach yields a promising result to enhance the accuracy of nursing diagnoses and improve work efficiency.**

Keywords—Nursing diagnosis process, nursing record documentation, natural language processing

I. Introduction

In recent years, deep learning technology has garnered significant attention for addressing various challenges in the healthcare field. A survey on karoshi (death from overwork) conducted in Japan revealed that the primary causes of excessive overtime among nurses were "emergency response and inpatient care" (73.6%) and "documentation, such as nursing records" (62.4%) [1]. Nursing records, which document the continuous nursing practices performed by nursing professionals in various situations [2], play a crucial role in the nursing process. Particularly, the formulation of nursing diagnoses is an essential process that guides the direction of nursing practice.

Automating nursing diagnoses holds a potential to reduce nurses' workloads and enhance the quality of nursing care. However, constructing a high-quality nursing diagnosis model requires a vast amount of nursing record text, presenting significant challenges due to the inclusion of patients' personal information in the records. Additionally, the lack of spaces between words in the Japanese language further complicates the issue.

Several text processing techniques are intensively introduced for analyzing nursing records to extract the text patterns and predict the trends from large volumes of text data to support nursing diagnoses. Techniques such as Bidirectional Encoder Representations from Transformers (BERT) [3] and Term Frequency-Inverse Document Frequency (TF-IDF) [4] are particularly useful for understanding the context of nursing records and extracting critical information.

This study proposes an architecture of nursing record analysis to assist nurses in the nursing diagnosis process. Under the preliminary experiments, the approach yields a promising result to enhance the accuracy of nursing diagnoses to reduce nurses' workloads and make a significant contribution to improving the quality of patient care.

Our contributions can be summarized as follows:

1) We propose a novel architecture utilizing cosine similarity to predict nursing diagnoses.
2) The Japanese version of BERT is employed to effectively handle and vectorize complex nursing record texts.

The remaining part of this paper is structured as follows. Section II reviews related work. Section III describes the nursing process. Section IV provides specific examples of the diagnosis part of the nursing process using simulated cases. In Section V, we propose an architecture of the nursing diagnosis process support model for nurses. Section VI presents the experimental setup and results, while Section VII discusses the findings. Finally, Section VIII concludes the paper.

II. Related Work

Recent advancements in deep learning have paved the way for significant improvements in patient care. Especially, NLP has been increasingly utilized to extract meaningful information from medical records. Lai et al. [5] applied NLP techniques to electronic health records (EHRs) to predict hospital readmission, achieving notable predictive performance. Another significant contribution by Cai et al. [6] involved using NLP for identifying patient cohorts for clinical trials by analyzing EHRs. These studies underscore the potential of NLP in automating the analysis of textual medical data, which is crucial for developing systems to support nursing diagnoses.

Focusing on the area of nursing, several studies have addressed the challenges of processing nursing records. Churpek et al. [7] demonstrated the feasibility of using machine learning techniques to predict patient deterioration from nursing records, supporting clinical decision-making using textual data. Furthermore, Li-Heng et al. [8] explored the use of deep learning models to classify nursing diagnoses from nursing notes, highlighting the challenges posed by unstructured text and the necessity for robust text processing techniques. These studies illustrate the complexity of nursing records and underscore the importance of sophisticated NLP methods to handle the nuances of medical terminology and documentation style.

In the context of enhancing nursing diagnoses through automated text analysis, the use of cosine similarity and logistic regression has shown significant promise. Li et al. (9) proposed a framework for supporting regularized logistic regression while ensuring data privacy, which is critical in

2024 IEEE International Symposium on Consumer Technology (ISCT) | 979-8-3503-6519-1/24/$31.00 ©2024 IEEE | DOI: 10.1109/ISCT62336.2024.10791266

handling sensitive patient information. Their method leverages distributed computing and cryptographic techniques to perform logistic regression without compromising data privacy, making it highly relevant for applications in the healthcare domain where patient confidentiality is paramount.

Li et al. [9] introduced an efficient optimization method based on distributed computing and strong cryptography to protect individual-level and summary data during logistic regression. This approach ensures that sensitive computations are partitioned and distributed, preventing the need to share raw individual data beyond the owning institutions.

In our study, we extend the principles from Li et al. [9] by applying cosine similarity and logistic regression to the analysis of nursing records. The nursing process involves detailed documentation of patient care, and automating the diagnosis process can significantly reduce the workload on nurses and improve the quality of patient care.

We calculate the cosine similarity between training cases and NANDA diagnosis vectors, utilizing these similarity scores as features for a logistic regression model to predict nursing diagnoses. This approach aligns with the methodologies discussed in Li et al. [9], leveraging the strengths of both cosine similarity for text vectorization and logistic regression for predictive modeling.

III. Nursing prosess

The nursing process, based on nursing knowledge and experience, serves as a systematic approach to providing an optimal and individualized care. The nursing process typically consists of five interrelated steps: assessment, nursing diagnosis, planning, implementation, and evaluation. These steps are interconnected and progress dynamically in a spiral manner, with the outcomes of each evaluation influencing subsequent assessments [10].

During the assessment phase, subjective and objective data are collected, organized, and analyzed. Subjective data include information directly obtained from the patient and/or his/her family through interviewing and external observation, expressed in the patient's own words, such as symptoms, health experiences, and concerns. For instance, the level of pain or the causes of anxiety fall under this category. Objective data consist of measurable or observable information collected through observations, physical examinations, and test results by nurses or other healthcare professionals. This includes physiological indicators such as blood pressure, body temperature, and heart rate, as well as test results and patient behaviors.

In the next step, nursing diagnosis, involves interpreting and analyzing the data gathered during the assessment. By interpreting and analyzing subjective and objective data based on knowledge and experience, nurses evaluate how the data relate to the patient's health issues and the consequent nursing needs. The problems identified through this interpretation and analysis form the basis of the nursing diagnosis, which helps determine the necessary nursing interventions and facilitates the planning process.

Frameworks such as Virginia Henderson's Nursing Need Theory (NNT) and Marjory Gordon's Functional Health Patterns (FHP) are utilized in the assessment process to comprehensively understand the patient and develop individualized care plans. Major systems for nursing diagnosis include the North American Nursing Diagnosis Association International (NANDA-I) [11] and FHP. Additionally, individual nurses may express diagnoses freely, or healthcare institutions may develop their own diagnostic frameworks.

In this study, we employ FHP for the assessment framework and NANDA-I for nursing diagnoses.

IV. Example of Nursing Process

Based on the aforementioned information, we will examine the actual implementation of the nursing diagnosis process for a male patient in his 50s who has undergone surgical treatment for gastric cancer.

Situation Setting

For about three months, the patient experienced epigastric discomfort after meals and took over-the-counter gastrointestinal medication. Approximately two weeks ago, he experienced severe upper abdominal pain and vomiting, prompting a visit to the outpatient clinic. Subsequent examinations, including gastroscopy, CT, and MRI, led to a diagnosis of gastric cancer (pyloric region cancer, Stage IIA, type L,2, T2, N1, H0, P0, CY0, M0), and he was admitted to the hospital for surgery. Preoperatively, the patient exhibited anxiety but actively participated in preoperative training, and the surgery was performed successfully. The procedure involved pylorus-preserving gastrectomy (2/3 resection) and lymph node dissection under general anesthesia.

Findings for the diagnosis of gastric cancer include the following primary tumor characteristics: (1) number and size of lesions, (2) affected area (three regions of the stomach: U (upper), M (middle), L (lower), (3) macroscopic classification, and (4) depth of invasion into the gastric wall (T). Additionally, information about metastasis is recorded, including (1) lymph node metastasis (N), (2) liver metastasis (H), (3) peritoneal metastasis (P), (4) peritoneal cytology (CY), and (5) distant metastasis (M). The stage of the cancer (Stage) is determined based on these findings.

For this patient, three nursing diagnoses were established on the second postoperative day: delayed postoperative recovery, ineffective coping, and bathing self-care deficit. Table 1 summarizes the data and assessments that led to these three nursing diagnoses. To explain the diagnostic process for delayed postoperative recovery, the assessments related to delayed postoperative recovery are highlighted in bold.

As shown in Table 1, the assessment of tissue damage from surgery, pain at the surgical site due to muscle contraction reflex during movement, and restricted oxygen intake led to the nursing diagnosis of delayed postoperative recovery. In a more comprehensive record, considering the postoperative context, additional evaluations would include assessing oxygenation through blood gas data such as PaCO2 and PaO2, as well as understanding the overall condition through blood tests.

V. Architecture of Nursing Diagnosis Process Support Model for Nurses

In the nursing records, data and text in subjective and objective parts can be grouped into 11 patterns, which can be labeled to the correspondence NANDA Knowledge base. Consequently, case similarity can be measured in terms of pattern similarity in details. On the other hand, NANDA Knowledge can be determined on the combination of patterns. In the process of searching for nursing assessment, we propose two similarity measures on the nursing records. The two primary methods can be considered for predicting nursing diagnoses by using the vectorized text data. In the first method, the Case-based Similarity involves integrating each pattern into a single case and calculating the similarity between cases. In the second method, NANDA-Oriented Similarity calculates the similarity for each pattern within a case. Both methods aim to identify patient characteristics and examine their similarity to nursing diagnoses.

In the Case-Based Similarity Calculation, the text volume tends to be larger as it treats each case as a whole, whereas NANDA-Oriented Similarity Calculation may result in a more limited text volume by focusing on individual patterns. Conversely, Case-Based Similarity Calculation allows for more detailed comparisons by evaluating each pattern individually (Figure 1).

For nursing diagnosis criteria, specific nursing diagnoses are matched based on existing NANDA labels. Another approach involves calculating the similarity based on the NANDA Knowledge Base (NANDA-Oriented Similarity Calculation) and predicting the most suitable diagnosis for new cases by assessing their correspondence to the nursing diagnosis labels.

Figure 1. How to calculate forecasts

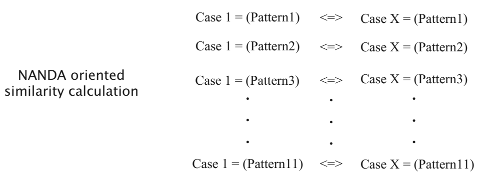

Table 1. Data and Assessment related to delayed postoperative recovery

<table>
<tr><th></th><th>Subjective data</th><th>Objective data</th><th>Assessment</th></tr>
<tr><td>Health Promotion</td><td>"I'd rather not use pain medication if possible. I'm fine as long as I stay still."</td><td>When encouraged to leave the bed post-surgery, the patient walks within the ward, but spends most other times lying in bed.</td><td>Postoperative mobilization occurs only when prompted, and the patient experiences significant incisional pain when moving, rarely initiating movement independently. The patient shows a passive approach to pain management. It is essential to enhance the patient’s understanding of the need for both mobilization and effective pain control.</td></tr>
<tr><td>Activity/ Rest</td><td>"I'd rather not use pain medication if possible. I'm fine as long as I stay still."</td><td>Anesthesia:
General anesthesia with epidural (Th7/8).
Surgical Procedure:
Pyloric-side subtotal gastrectomy plus lymph node dissection.
Reconstruction Method:
Billroth I procedure.
Surgical Duration: 3 hours.
Intraoperative Fluids: 2,400 ml; Blood Loss: 200 ml.
Abdominal Drain Placement.
Oxygen Administration: [Immediately post-surgery]
O2 at 4L, 28% via face mask
[Day 1-2 post-surgery]
Nasal cannula at 2L]
Respiration, SpO2 (oxygen saturation), blood pressure, pulse rate.
<table>
<tr><th></th><th>Pre-operative</th><th>Post-operative</th><th>Day 1</th><th>Day 2</th></tr>
<tr><td>Respiration rate(/min)</td><td>14</td><td>16-20</td><td>14-18</td><td>14-18</td></tr>
<tr><td>S_pO_2(%)</td><td></td><td>98-99</td><td>98-99</td><td>99</td></tr>
<tr><td>Blood pressure (mmhg)</td><td>120/80</td><td>110-128/ 76-84</td><td>112-120/ 76-82</td><td>110-118/ 74-80</td></tr>
<tr><td>pulse rate</td><td>68</td><td>76-92</td><td>74-84</td><td>72-80</td></tr>
</table>
No wheezing, arrhythmia, or lung sounds noted.
No cyanosis.
Productive cough with white sputum
Continuous administration of analgesics via epidural catheter and intravenous drip. Indomethacin suppository 50mg administered during episodes of increased pain.
Day 1: Assisted ambulation within the ward, supported by an IV pole and nurse.
Day 2: Assisted ambulation to the toilet and washroom, supported by an IV pole.</td><td>Pain post-surgery is attributed to tissue damage and the muscle spasm reflex. Continuous administration of analgesics is being conducted via an epidural catheter and an intravenous drip. For episodes of increased pain, an indomethacin suppository (50mg) is administered. However, the patient experiences significant pain during movement, leading to minimal activity, primarily prompted by necessity. Additionally, the patient has an ongoing intravenous drip in the left forearm and an abdominal drain, which restrict mobility during walking. Consequently, the patient often remains in a recumbent position, which may inhibit adequate airway clearance. The reduction in physical activity could lead to insufficient oxygen uptake and potential blood pooling due to a loss of vascular tone, contributing to decreased venous return. Therefore, there is a high possibility of delayed postoperative recovery.</td></tr>
<tr><td>Perception/ Cognition</td><td>"I have trouble sleeping at night after getting up to use the bathroom several times. The movement also makes the incision hurt. Even when I wake up in the morning, I don't feel well-rested."</td><td>Epidural Tube: 0.25% Marcaine injection.
Intravenous Drip: Continuous administration of Fentanyl 0.1 mg × 30 + Normal Saline (100) × 1.8.
During the action of getting out of bed, there is significant pain at the incision site, and the patient slowly rises while holding onto the bed rails. When moving to the toilet, the patient walks with a forward-leaning posture while holding the abdomen with one hand.
Coughing does not effectively engage the abdominal muscles, making it difficult to expectorate.</td><td>Post-surgery, pain at the incision site has developed due to tissue damage and muscle spasm reflex.
Continuous administration of analgesics via the epidural tube and IV drip is in place, and suppositories are used during times of increased pain. However, based on facial expressions and verbal expressions, it is suspected that adequate pain control has not been achieved. Therefore, it is necessary to closely monitor the pain's status and characteristics and adjust the management accordingly.</td></tr>
<tr><td>Coping/ Stress tolerance</td><td>"I'm in pain and can't move my body the way I want to," he said, visibly frustrated.</td><td>He is meticulous and strict in his professional life. Recently, he was transferred to a different department.
When his wife and children visited, he frequently exhibited irritable behavior.</td><td>Postoperatively, the patient is unable to move freely due to pain during movement, resulting in frustration and inadequate pain management. Consequently, his activity level has decreased, and mobilization is delayed, which could potentially lead to complications such as delayed intestinal recovery, anastomotic failure, and pulmonary complications. It is considered necessary to improve pain management to enable better coping mechanisms.</td></tr>
</table>

VI. EXPERIMENT

A. Abbreviations and Acronyms

A literature search was conducted through the National Diet Library (NDL) ONLINE, focusing on publications from 2018 to 2023 with "nursing process" in the title. Works related to "psychiatric," "maternal," "pediatric," or "home care" nursing were systematically excluded to narrow the selection to publications in digital format, facilitating access and data set compilation. Consequently, a data set was created from four books containing case studies suitable for extracting data and information for nursing diagnose. From these four books, 29 case studies were extracted.

B. Data Preprocessing

In this study, we aimed to predict nursing diagnoses by vectorizing text using BERT. The following sections detail the data preprocessing steps.

First, data cleaning involved the removal of unwanted characters and the standardization of text data. Specifically, characters within full-width parentheses were removed, and non-Japanese characters were excluded to clean the text.

Next, for each case classified into Gordon's 11 patterns, we conducted analysis using two methods as shown in Figure 1. The first method, referred to as Oriented Text, involved combining the 11 patterns into a single text and vectorizing it. The second method, referred to as Pattern Text, involved vectorizing each pattern individually and then averaging them to create a vector representing the case.

Text vectorization was performed using the Japanese version of the BERT model. After cleaning the text for each pattern, the BERT model was used to convert the text into vectors, resulting in vectors for each pattern of each case. These vectors were then averaged to create a single vector representing each case.

The dataset was split into 28 cases for the training set and the remaining case for the test set. Additionally, texts related to each NANDA-I diagnosis label (including definitions, diagnostic indicators, risk factors, related factors, high-risk groups, and related conditions) were vectorized. The average vectors of these texts were calculated to generate comprehensive vectors representing the characteristics of each diagnosis.

C. Training and Prediction Details

Cosine similarity between the vectors of each training case and the NANDA diagnosis vectors is calculated. The nursing diagnoses (labels) for the training cases are based on the corresponding NANDA diagnosis labels. For example, if Training Case 1 has the labels "Acute Pain" and "Anxiety," the cosine similarity between the vector of this case and the NANDA diagnosis vectors for "Acute Pain" and "Anxiety" is calculated. The calculated cosine similarity scores are used as features, providing similarity scores for all NANDA diagnosis labels for each training case. This results in a feature vector of NANDA diagnosis similarity scores for each training case.

A logistic regression model is then used to learn from these similarity scores and the nursing diagnosis labels of the training cases. The model is trained to use the similarity scores for each NANDA diagnosis label to predict the corresponding nursing diagnosis labels for each training case.

For the test case, the same process is applied: the test case is vectorized using BERT, and the cosine similarity between its vector and the NANDA diagnosis vectors is calculated. This provides similarity scores for all NANDA diagnosis labels for the test case. The trained logistic regression model is then used to predict the nursing diagnoses.
for the test case based on these similarity scores. The model uses the similarity scores of the test case to predict the most appropriate nursing diagnoses.

D. Evaluation

The accuracy of the predictions was calculated using several steps. First, the true diagnoses, which are the actual correct labels for the test cases, were identified. Next, the predicted diagnoses, or the labels predicted by the model, were determined. The true positives (TP), which represent the number of correct labels that were also predicted correctly, were then counted. Additionally, the false positives (FP), which are the number of incorrect labels that were predicted, and the false negatives (FN), which are the number of correct labels that were not predicted, were identified.

The formula for accuracy is:

$$\text{Accuracy} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \tag{1}$$

This formula indicates the proportion of correct labels that were correctly predicted out of all actual correct labels. False Positives are not considered in this accuracy calculation.

E. Evaluation Results

In this study, we conducted experiments using 29 different patterns by splitting the dataset into 28 training sets and 1 test case each time. The resulting accuracies ranged from 0 to 0.67. Here, we present the results of the case that showed the best outcome in Table 2 and focus on discussing these results.

Table 2. The results of best outcome

Ground Truth	Oriented Text		Pattern Text	
	Diagnosis	*Degree of similarity*	*Diagnosis*	*Degree of similarity*
Impaired Comfort	**Anxiety**	**0.404**	**Anxiety**	**0.428**
Anxiety	Bathing self-care deficit	0.403	Bathing self-care deficit	0.393
Ineffective health self-management	Risk for adult falls	0.255	**Impaired Comfort**	**0.250**
	Impaired Comfort	**0.238**	Risk for adult falls	0.215
	Toileting self-care deficit	0.231	Readiness for enhanced self-care	0.214
	Readiness for enhanced self-care	0.223	Toileting self-care deficit	0.214
True Positive	2		2	
False Positive	4		4	
False Negative	1		1	
Accuracy	0.667		0.667	

In each experiment, 28 cases were used as the training set, and the remaining case was used as the test set. For each case, text vectorization was performed using two methods: Oriented Text and Pattern Text, and nursing diagnoses were predicted using each method.

First, logistic regression models were trained using similarity scores for each diagnosis label to predict the diagnoses of the test case. The accuracy of the predictions was evaluated using the metrics of TP, FP, FN, and Accuracy.

In the case that showed the highest accuracy, the following prediction results and actual diagnosis results were obtained. The model's predictions for this case using Oriented Text were: Anxiety (0.404), Bathing self-care deficit (0.403), Risk for adult falls (0.255), Impaired Comfort (0.238), Toileting self-care deficit (0.231), and Readiness for enhanced self-care (0.223).

From these results, the predicted diagnoses that matched the actual diagnoses were Impaired Comfort and Anxiety. The True Positives (TP) were 2, FP were 4, FN were 1, and the overall accuracy was 0.667. Using Pattern Text, the degree of similarity for each diagnosis slightly improved, but there were no changes in other evaluation metrics.

VII. Discussion

The BERT-based vectorization and logistic regression model used in this study demonstrated high accuracy in predicting certain nursing diagnoses. However, challenges remain in improving overall accuracy. Healthcare institutions record information in various formats, and machine learning has played a crucial role in medical research. However, concerns regarding patient privacy limit the broader machine learning research community's access to patient data, resulting in slower progress compared to other fields [12]. This study utilized a limited number of simulated cases, but improving model accuracy requires the collection of more diverse case data. In particular, improving the balance of each diagnosis label could enhance the model's learning effectiveness. Therefore, the collection of large datasets and the anonymization of patient data are necessary.

One potential solution to this challenge is the utilization of large language models (LLMs), which can augment medical data and generate high-quality data despite a lack of volume and diversity [13]. Data augmentation is considered effective in compensating for data scarcity and maintaining diversity, thereby strengthening model generalization.

Moreover, while BERT was used for vectorization in this study, it is also necessary to consider the introduction of more advanced text analysis techniques and other NLP models. For instance, the use of pre-trained models specialized in the medical field should be considered [14]. This could allow for a more accurate capture of the characteristics of Japanese medical records and improve diagnostic accuracy.

These efforts are expected to build more accurate nursing diagnosis prediction models, enhancing their practicality in clinical settings. Future research will aim to develop even higher-accuracy prediction models by incorporating additional data and new models.

VIII. Conclusion

The performance of the BERT model demonstrated its potential in predicting nursing diagnoses. However, further improvements and larger datasets are necessary for more robust predictions. This study demonstrated the feasibility of analyzing nursing records and predicting nursing diagnoses using text vectorization techniques.

One significant challenge is the limited dataset size, which poses a major obstacle in evaluating the effectiveness of the proposed system. Increasing the dataset size is crucial for enhancing the system's predictive accuracy. Collecting and utilizing larger datasets will enable more reliable and robust model predictions. Future challenges include expanding the dataset and improving the model's accuracy by exploring additional machine learning techniques to further enhance predictive performance.

References

[1] Ministry of Health, Labour and Welfare. White Paper on Measures to Prevent Karoshi (Overwork Death) (Annual Report for Fiscal Year 2017) [Internet]. Available from: https://www.mhlw.go.jp/wp/hakusyo/karoushi/18/index.html

[2] Japanese Nursing Association. Guidelines for Nursing Records [Internet]. 2018. Available from: https://www.nurse.or.jp/home/publication/pdf/guideline/nursing_record.pdf

[3] Devlin J, Chang M-W, Lee K, Toutanova K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding [Internet]. arXiv [cs.CL]. 2018. Available from: http://arxiv.org/abs/1810.04805

[4] Juan R. *Using TF-IDF to Determine Word Relevance in Document Queries*. Proceedings of the first instructional conference on machine learning. 2003;242(1).

[5] Lai H-J, Tan T-H, Lin C-S, Chen Y-F, Lin H-H. Designing a clinical decision support system to predict readmissions for patients admitted with all-cause conditions. J Ambient Intell Humaniz Comput [Internet]. 2020 Jan 2; Available from: http://dx.doi.org/10.1007/s12652-019-01579-6

[6] Cai T, Giannopoulos AA, Yu S, Kelil T, Ripley B, Kumamaru KK, et al. Natural language processing technologies in radiology research and clinical applications. Radiographics. 2016 Jan;36(1):176–91.

[7] Churpek MM, Yuen TC, Winslow C, Meltzer DO, Kattan MW, Edelson DP. Multicenter comparison of machine learning methods and conventional regression for predicting clinical deterioration on the wards. Crit Care Med. 2016 Feb;44(2):368–74.

[8] Li-Heng F, Chris K, Kenrick C, Adler P, Kang M-J, Patricia C D, et al. *Utilizing timestamps of longitudinal electronic health record data to classify clinical deterioration events*. J Am Med Inform Assoc. 2021;28((9)):1955–63.

[9] Li W, Liu H, Yang P, Xie W. Supporting regularized logistic regression privately and efficiently. PLoS One. 2016 Jun 6;11(6):e0156479.

[10] American Nurses Association. Nursing: Scope and Standards of Practice. Nursesbooks; 2021. (4th Edition).

[11] Herdman, Heather T, Kamitsuru S, Lopes CT. Nanda International Nursing Diagnoses 12th ed. Thieme Medical Pub; 2021.

[12] Kaissis GA, Makowski MR, Rückert D, Braren RF. Secure, privacy-preserving and federated machine learning in medical imaging. Nat Mach Intell. 2020 Jun 8;2(6):305–11.

[13] OpenAI. GPT-4 Technical Report. 2023.

[14] Huang K, Altosaar J, Ranganath R. ClinicalBERT: Modeling clinical notes and predicting hospital readmission [Internet]. arXiv [cs.CL]. 2019. Available from: http://arxiv.org/abs/1904.05342