ECTI Transactions on Computer and Information Technology
Journal homepage: https://ph01.tci-thaijo.org/index.php/ecticit/
Published by the ECTI Association, Thailand, ISSN: 2286-9131

Assessment Pattern Mapping in NANDA-I Nursing Diagnoses Framework by BERT Approach

Kubo Takahiro[1], Virach Sornlertlamvanich[2] and Thatsanee Charoenporn[3]

ABSTRACT

Nurses play a crucial role in healthcare, directly influencing patient care quality. With a global nursing shortage, enhancing nursing efficiency and care quality is urgently needed. This foundational study explores the advantages of text and data processing techniques to determine NANDA-I nursing diagnoses using both subjective and objective patient data recorded by nurses. By employing text data similarity analysis and a prototype of the predictive model, our research aims to refine the nursing assessment process and facilitate the automation of nursing diagnoses. This work highlights the accuracy of BERT-based assessment pattern matching to support nursing practices and sets a platform for future research to address the nursing shortage effectively.

Article information:

Keywords: Nursing Diagnosis, Prediction, Degree of Similarity

Article history:
Received: July 18, 2024
Revised: September 19, 2024
Accepted: October 10, 2024
Published: November 30, 2024 (Online)

DOI: 10.37936/ecti-cit.2025191.257350

1. INTRODUCTION

Nurses constitute one of the most significant human resources in the healthcare sector, playing a pivotal role in determining the quality of patient care. According to the World Health Organization's (WHO) 2020 report, the global shortage of nurses is estimated to be 5.9 million [1], and in Japan, a potential shortage of up to 250,000 nurses is projected by 2025 [2]. This presents a significant challenge for Japan's healthcare industry, which faces urgent needs for cost reduction and operational efficiency amidst an ageing society and the insufficiency of professionals. In recent years, deep learning technologies have garnered significant attention for solving numerous challenges in the healthcare sector. Notably, a survey conducted in Japan regarding death from overwork indicated that the primary causes of excessive overtime among nurses were "emergency response to emergencies and inpatient care" (73.6%), followed by "documenting nursing records and other paperwork" (62.4%) [3].

Against this background, researchers believe that the utilization of artificial intelligence (AI) in nursing has the potential not only to improve operational efficiency but also to directly contribute to securing nursing time and improving the quality of nursing.

Nursing records, documented daily by nurses, are vital documents recording observations and judgments acquired through daily care. Mitha et al. conducted a review study on the potential of applying Natural Language Processing (NLP) to nursing records for predicting nursing diagnoses and supporting clinical decision-making [4]. However, insights into the applicability of NLP in nursing remained limited. This highlights the underutilization of nursing records in improving nursing practice.

The automation of nursing diagnoses holds the potential to alleviate the workload of nurses and enhance the quality of nursing care. However, constructing accurate nursing diagnosis models necessitates large-scale text datasets. Privacy concerns related to the personal information in nursing records hinder data acquisition. Moreover, the characteristic of languages like Japanese, which do not employ spacing between words, further complicates this issue.

In this study, we propose a model to determine NANDA International (NANDA-I) (2021-2023) [5] nursing diagnoses from data collected on patients during assessment. Through this research, we aim to explore methods for predicting nursing diagnoses using NLP techniques. Our objective is to contribute to

[1]The author is with the Factory of Nursing, Faculty of Data Science, Doctoral degree, Musashino University, Japan, Email: ta_kubo@musashino-u.ac.jp
[2,3]The authors are with the Asia AI Institute (AAII), Faculty of Data Science, Musashino University, Japan, E-mail: virach@musashino-u.ac.jp and thatsanee@ds.musashino-u.ac.jp
[2]The author is with the Faculty of Informatics, Burapha University, Thailand, E-mail: virach@musashino-u.ac.jp
[1]Corresponding author: ta_kubo@musashino-u.ac.jp

enhancing decision-making processes in nursing practice, thereby ultimately improving the quality of patient care.

The structure of this paper is as follows: Section 2 organizes the thought processes for nursing practice and explains the framework practically used. Section 3 reviews the efforts of applying NLP techniques in the nursing diagnostic process. Section 4 describes the proposed method. Section 5 presents the experiment results, and Section 6 discusses the results and prospects. Finally, Section 7 concludes the paper.

2. THOUGHT PROCESS OF NURSES IN NURSING ASSESSMENT

In Japan, nursing involves licensed nursing professionals (including midwives, public health nurses, registered nurses, and licensed practical nurses) practicing across various settings in health, medical, and welfare sectors, targeting individuals, families, groups, and community societies of all ages [6]. This study focuses on registered nurses working in hospital wards, where a majority of nursing activities are concentrated.

The nursing process, regarded as a systematic method for providing optimal and individualized care based on nursing knowledge and experience, is critical for assessing health issues and delivering appropriate nursing interventions. It typically encompasses five steps: assessment, nursing diagnosis, planning, implementation, and evaluation. These steps are interrelated and dynamically progress spirally, where evaluation informs the subsequent assessment [7, 8], as shown in Fig. 1.

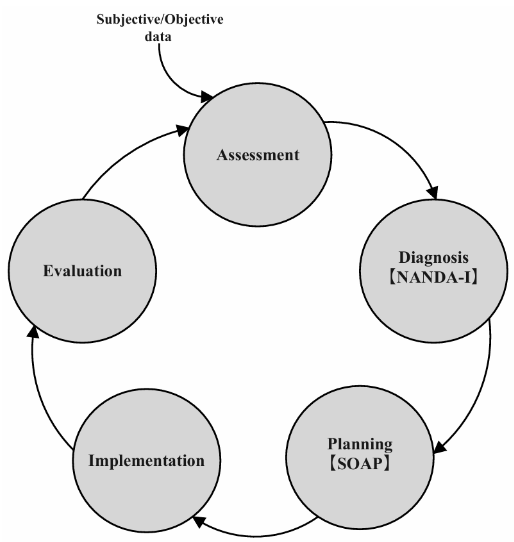


Fig.1: *Five Steps of The Nursing Process.*

Each step incorporates theories or frameworks to develop individualized care plans. In this study, we utilize Marjory Gordon's Functional Health Patterns (MGFHP) during the assessment phase. The MGFHP consists of 11 comprehensive patterns for evaluating a patient's overall health status. Subsequently, we apply NANDA-I, the nursing diagnosis framework established by the North American Nursing Diagnosis Association, during the diagnosis phase. NANDA-I is a structured framework that identifies patient health problems and risks, facilitates the planning of effective nursing interventions, and supports evidence-based nursing practice. It includes diagnostic labels, definitions, related factors, and defining characteristics.

The nursing process begins with the assessment phase. This phase involves collecting and organizing data about the patient, including health-related and social background information, to capture a holistic view of the patient. Several frameworks exist for capturing patient data; this study employs MGFHP, which are widely used by nurses in Japanese hospitals for data collection. Following the phase of assessment is the nursing diagnosis, where multiple necessary nursing diagnoses are derived based on the assessed data and information, and the required nursing diagnosis for the patient is determined. Nursing diagnoses are either identified independently by nurses or derived from frameworks such as NANDA-I. Following the diagnosis, a nursing plan is created, which informs subsequent nursing practice and evaluation. Once a nursing diagnosis is established, a nursing plan based on the diagnosis is formulated, leading to nursing practice and evaluation. The Nursing Interventions Classification (NIC) [9] and Nursing Outcomes Classification (NOC) [10] systems classify nursing practices and outcomes. These systems are designed to ensure consistent care quality throughout the nursing process. The thought process leading to nursing practice is organized by the nursing process, where nurses interpret and analyze the assessed data and information to identify nursing problems unique to each patient. In Japan, most of this process is carried out by the nurses themselves, and inappropriate nursing diagnoses can lead to failure to provide appropriate care, to the detriment of the patient.

Nurses document these planning processes in nursing records, often using the SOAP format, which stands for Subjective data, Objective data, Assessment, and Plan. Subjective data are the patient's statements, gathered through personal interviews, while objective data include observed facts and measurements. Nurses use their expertise to interpret and analyze both subjective and objective data. This helps them identify the patient's needs and health issues. The plan is based on the assessment, with the nurse formulating a nursing diagnosis to develop a care plan. This plan includes future actions and interventions tailored to the patient's unique needs, aiming to achieve the desired health outcomes.

This study, therefore, proposes the development of a system to support the process from assessment to planning. Using NANDA-I as the foundation for nursing diagnoses provides a structured framework for predicting nursing diagnoses from collected assessment data. The construction of this system will enable a consistent connection to NIC and NOC from collected data. It is anticipated that even nurses with less experience or those working in new areas can formulate appropriate nursing diagnoses.

3. EFFORTS OF APPLYING NLP TECHNIQUES IN THE NURSING DIAGNOSTIC PROCESS

A search of the Medical Journal Web for previous studies in Japan utilizing NLP techniques in the field of nursing revealed the following studies: a study on the extraction of pain expressions from medical records using BERT [11], a study that attempted to extract time series knowledge about the course of medical conditions and treatment effects by text mining from discharge summaries, which are medical documents [12], and a study® that used Studio® to text-mine the content of interviews with nurses to verify the effectiveness of a nursing delivery method [13].

Internationally, 43 review studies using NLP for nursing records highlighted the relatively small but growing number of publications on NLP in nursing and the potential for NLP to expand the methods and findings in the future through appropriate performance measures and existing standard nursing terminology [14]. Other studies used NLP to identify communication failures between home care nurses and physicians and to evaluate communication [15], and a study on creating an NLP algorithm to extract wound infection-related information from nursing records [16].

While the performance of computers has dramatically improved, coupled with the promotion of research methods based on AI, and the number of studies incorporating data science techniques in the nursing field is increasing internationally, the number of such studies in Japan is extremely small [17–19]. In addition, as far as we could find, there were no studies on the nursing diagnostic process conducted by nurses.

These considerations highlight the transformative potential of NLP in the nursing field, ranging from extracting critical patient information from unstructured data sources. Integrating these technologies into nursing represents an important step toward achieving better healthcare outcomes and greater efficiency. The uniqueness of this paper lies in its application of natural language processing (NLP) techniques to enhance the accuracy and efficiency of nursing diagnosis. While previous research has focused on the analysis of general medical documents, the use of NLP in the specific domain of nursing diagnosis has not yet been fully explored. Our approach is to explore the potential value of textual data routinely recorded by nurses and to directly link the results to improved accuracy in nursing diagnosis and quality of patient care.

4. ASSESSMENT PATTERN MAPPING FOR NANDA-I DIAGNOSES

Nursing records are comprehensive, chronicling every aspect of a patient's hospitalization experience. This study harnesses frameworks like MGFHP for systematic categorization. Despite this, it is recognized that nursing records may also encompass routine details such as gender and weight. The overview of case processing is depicted in Fig. 2.

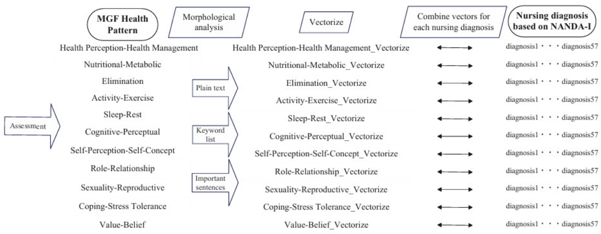


Fig.2: *Overview of Case Processing.*

According to MGFHP, the subjective and objective data align with the 11 patterns. Accordingly, we have adopted three distinct text-processing methodologies to ascertain their influence on the predictive accuracy of NANDA-I diagnoses.

For this study, we focused on evaluating the performance of two models, BERT and SBERT, in combination with each of the three text processing methods applied. The goal was to determine which combination of model and pre-processing method would yield the best predictive accuracy. We assessed whether BERT and SBERT show significant performance differences when paired with plain text, keyword lists, or important sentences, and whether certain combinations enhance the model's ability to predict accurate nursing diagnoses.

The first method, "plain text", operates on the premise that all nurse-documented information is imperative to capture a comprehensive patient profile. It treats subjective and objective data as a singular textual continuum, preventing fragmentation and maintaining the textual context.

For the second method, "keyword list", morphological analysis was performed on the plain text, and after excluding numerals, nouns, verbs, adjectives, shape verbs, and pronouns were extracted. After extraction, the value of Term Frequency-Inverse Document Frequency (TF-IDF) calculation was applied to extract the ten most salient keywords. This method assumes that extracting nursing diagnoses can be facilitated by focusing on critical elements in the documented patterns.

The final method, known as "important sentences," extracts sentences from plain text that incorporate the critical terms identified in the keyword list. This approach was informed by Sitikhu *et al.*, who demonstrated that cosine similarity, when paired with TF-IDF vectors, is notably productive in detecting similarities amongst brief news articles.

This study used several NLP techniques to improve the predictive accuracy of NANDA-I nursing diagnoses. Specifically, morphological analysis was first used to extract meaningful information from patient records to identify key nouns, verbs, and adjectives. We then calculated the TF-IDF from these words and used cosine similarity to assess the similarity between texts. This laid the foundation for our analysis by quantifying textual features to predict relevant nursing diagnoses.

4.1 Datasets

To support this study, Table 1 lists the four selected books, chosen based on refined search criteria for their relevance to case studies that allow for the extraction of data necessary for nursing diagnosis derivation. From these four books, 29 cases were extracted, as shown in Table 2.

Table 1: *Target Books.*

No	Title	Publication year	Author	Publisher	Number of cases
1	Disease-Specific Nursing Process Seminar - Volume 1 [疾患別看護過程セミナー上巻]	2021	Yukihiro Yamada	Scio Publishing	11
2	Disease-Specific Nursing Process Seminar - Volume 2 [疾患別看護過程セミナー下巻]	2021	Yukihiro Yamada	Scio Publishing	7
3	Guide to the Development of the Nursing Process: How to Write a Practice Record, Volume 2 [実習記録のかきかたがわかる 看護過程展開ガイド 第2版]	2022	Kazuko Ren	Shorin-company	8
4	Nursing Process Development Guide by Area [領域別 看護過程 展開ガイド]	2022	Kazuko Ren	Shorin-company	3

A literature search was conducted using the National Diet Library (NDL) ONLINE, focusing on publications between 2018 and 2023 that included "nursing process" in their titles. To narrow the scope, works related to "psychiatric," "maternal," "pediatric," or "home care" nursing were systematically excluded. This refinement ensured a focus on digital format publications, facilitating easy access and dataset compilation.

Within this defined period, 61 publications were identified that initially met the "nursing process" search criteria. However, further exclusions were applied based on predefined criteria: 22 titles were excluded for containing specific keywords, 17 for lacking digital availability, and 18 for not including case-based Subjective and Objective data. Ultimately, the dataset was compiled from the four books in Table 1, each containing case studies suitable for extracting the data necessary for nursing diagnosis derivation.

4.2 Data collection

In the analysis of 29 case studies derived from four books, the following methodological steps were undertaken:

Table 2: *Case list.*

Case No	Book No	Disease
1	3	Retinal detachment
2	4	Gastric cancer
3	4	Chronic obstructive pulmonary disease
4	4	Sequelae of cerebral haemorrhage
5	3	Cerebral infarction
6	3	Acute myocardial infarction
7	3	Multiple sclerosis
8	1	Lung cancer
9	1	Heart failure
10	1	Angina pectoris
11	1	Myocardial infarction
12	1	Gastric cancer
13	1	Colorectal cancer
14	1	Liver cirrhosis
15	1	Subarachnoid haemorrhage
16	1	Cerebral infarction
17	2	Renal failure
18	2	Benign prostatic hyperplasia
19	2	Femoral neck fracture
20	2	Parkinson's disease
21	2	Diabetes mellitus
22	2	Dementia
23	2	Breast cancer
24	3	Gastric cancer
25	3	Lung cancer
26	3	Breast cancer
27	3	Diabetes mellitus
28	1	Pneumonia
29	1	COPD

1. Extraction of Subjective data, Objective data, and nursing diagnoses from each case study.
2. The assessment framework employed for Subjective and Objective data categorization was based on MGFHP. This framework includes 11 patterns, namely Health Perception-Health Management, Nutritional-Metabolic, Elimination, Activity-Exercise, Sleep-Rest, Cognitive-Perceptual, Self-Perception-Self-Concept, Role-Relationship, Sexuality-Reproductive, Coping-Stress Tolerance, and Value-Belief. Cases already aligned with Gordon's framework were directly classified accordingly, whereas cases utilizing alternate frameworks underwent manual classification by the researcher to fit Gordon's categories.
3. For cases where explicit NANDA-I nursing diagnosis labels were absent, each case underwent a detailed review to ensure that the most appropriate NANDA-I label was manually identified and assigned. A particular case was excluded from this study due to the absence of an applicable NANDA-I counterpart.
4. Subsequently, the study represented 56 utilized NANDA-I nursing diagnosis labels in binary form for each analysed case.

4.3 Experiments

This research employed text vector representations, contextualized using the BERT model, to analyse plain text, keyword lists, and important sentences. The utilized BERT model, pre-trained on Japanese text [22], alongside a Sentence-BERT (SBERT) model [23] designed explicitly for inter-sentence similarity assessments, draws upon their demonstrated success across a spectrum of Natural Language Processing (NLP) tasks involving Japanese text. Kawazoe et al. developed a clinical-specific BERT model employing a substantial corpus of Japanese clinical texts, showcasing its superior performance in medical text analytics compared to models pre-trained on generic domains [24]. SBERT, an adaptation of the original BERT architecture, facilitates direct optimization for similarity comparisons [23], rendering it exceptionally apt for analysing nursing issue narratives and discerning their relationships with textual depictions.

A matrix was formulated, intertwining nursing diagnoses with corresponding feature vectors derived from this vectorized textual data. Through Cosine similarity computations, the research assessed the unity between the feature vector of a test case and those within the dataset, facilitating the prognostication of the nursing diagnosis most akin to the test case. Furthermore, a matrix amalgamating the feature vector X (integrating the 11 subjective and objective text patterns from Gordon's Functional Health Patterns) with the binary-encoded Y (NANDA-I nursing diagnosis labels) was established. This similarity-driven comparative analysis evaluated the vectorized representation of textual data, identified the most similar cases within the dataset, and enabled the prediction of NANDA-I nursing diagnosis labels.

The linkage between the feature vector X (merging the 11 distinct functional health patterns per textual case) and Y (nursing diagnosis) was vectorized, reflecting the array of functional health patterns unique to each text case. These feature vectors, coupled with the binary-coded nursing issues (NANDA-I labels), formed the basis for the matrix construction.

In the predictive phase, the research first generated average feature vectors for each nursing issue from 28 cases, excluding a singular test case from 29. Cosine similarity was then calculated between the feature vector of the test case and the mean feature vector corresponding to each nursing issue, enabling a comparison of their respective similarities. The average vector computation and cosine similarity calculation are pivotal in this study for quantitatively elucidating the nexus between textual descriptions and nursing issues. By aggregating feature vectors into an average vector for each nursing issue, the method encapsulates the textual features' central tendencies associated with each issue. This procedure mirrors the creation of representative signatures for each nursing issue, simplifying nuanced comparisons between the test case and predefined categories.

Cosine similarity acts as a measure to quantify the alignment between the vector of the test case and the average vectors for each nursing issue. Thanks to its normalization feature, this metric permits text comparisons based on directional orientation within the feature space rather than size, enabling more semantically concentrated analyses. This strategy, corroborated by Sun et al. through their exploration of BERT's fine-tuning for text classification, underscores the adaptability and effectiveness of BERT embeddings in capturing textual subtleties [25].

Following the outlined methodology, the study organized nursing issues with the highest five and lowest three similarity ratings into a table, derived from the conducted calculations.

5. EXPERIMENT RESULTS

Upon implementing the proposed methodology, our experiments consistently identified relevant nursing problems across all evaluation patterns. This section examines the results for test cases 4 and 26.

The tables summarize the predicted top 5 and bottom 3 NANDA-I nursing diagnoses labels and their corresponding NANDA-I numbers (Table 3). It should be noted that items in bold within the tables indicate a match with the ground truth, and italicized items signify that the similarity threshold exceeds 0.75.

Table 3: *NANDA-I No and NANDA-I Diagnosis Labels.*

NANDA-I No	NANDA-I diagnosis labels
002	Imbalanced Nutrition: Less Than Body Requirements
007	Hypothermia
011	Constipation
027	Deficient fluid volume
028	Risk for imbalanced fluid volume
039	Risk for Aspiration
047	Risk for impaired skin integrity
064	Parental Role Conflict
083	Decisional conflict
093	Fatigue
100	Delayed surgical recovery
102	Feeding self-care deficit
108	Bathing self-care deficit
109	Readiness for enhanced self-care
110	Toileting self-care deficit
125	powerlessness
129	Chronic confusion
132	Acute pain
146	Anxiety
147	Death anxiety
152	Risk or powerlessness
214	Impaired Comfort
247	Risk for impaired oral mucous membrane integrity
266	Risk for surgical site infection

267	Risk for unstable blood glucose level
276	Ineffective health self-management
281	Risk for ineffective lymphedema self-management
297	Disability- associated urinary incontinence
299	Risk for decreased activity tolerance
303	Risk for adult falls
303	Risk for adult falls
304	Risk for adult pressure injury

To make it easier to understand how the models process nursing records and generate predictions, we will show an example of a patient's input record and the corresponding outputs at various stages: plain text, keyword list, and important sentences. The following Table 4 illustrates the pattern for Health Perception-Health Management in case 1, demonstrating how the SBERT model processes the data before generating predictions.

Table 4: *Target Books.*

Plain text	Since last summer, I haven't had an appetite, and I feel lethargic. I'm not sure what I should do moving forward. I have no idea where to go from here. A 57-year-old woman, diagnosed with gastric cancer (stage IIA). Since last summer, I have experienced a loss of appetite, and since autumn, nausea and *belching* have appeared. After undergoing a detailed examination at *Hospital* A, I was diagnosed, and I was referred to General *Hospital* B for surgery. The current plan is to perform a total gastrectomy (under general *anaesthesia* and *epidural anaesthesia*), followed by reconstruction using the *Roux-en-Y* method. My past medical history includes appendicitis at the age of 11. I am not currently taking any medications. Height: 160.0 cm, *weight*: 55 *kg* (a 10 *kg weight* loss over the past *six months*). I do not smoke or drink alcohol, and I have no allergies.
Keyword list	Anaesthesia, kg, weight, hospital, diagnosis, six months, epidural, Roux-en-Y, belching
Important sentences	I have no idea where to go from here. A 57-year-old woman, diagnosed with gastric cancer (stage IIA). Since last summer, I have experienced a loss of appetite, and since autumn, nausea and belching have appeared. After undergoing a detailed examination at Hospital A, I was diagnosed and referred to General Hospital B for surgery. The current plan is to perform a total gastrectomy (under general anesthesia and epidural anesthesia), followed by reconstruction using the Roux-en-Y method. Height: 160.0 cm, weight: 55 kg (a 10 kg weight loss over the past six months)

Remark: Underscored text indicates important sentences; Italic text indicates keywords.

Initially, the BERT model's results for case 4 are presented in Table 5, while the SBERT model's results are detailed in Table 6.

The BERT analysis for case 4 accurately identified two nursing issues as the ground truth, placing them in the top two positions across plain text, keyword lists, and important sentences, with remarkably high similarity scores above 0.9. Nonetheless, it was observed that similarity scores for positions three to five, which did not align with correct answers, remained steadfastly above 0.8. Furthermore, the three lowest-ranked nursing issues demonstrated similarity scores exceeding 0.7.

In the case of the SBERT-derived results for case 4, the model also predicted two nursing issues as the ground truth within the top two ranks across plain text, keyword list, and important sentences. However, the similarity scores slightly decreased, remaining just above 0.8, compared to the BERT experiment. Conversely, for ranks three to five, which did not match the correct answers, similarity scores were recorded at 0.6 for plain text, keyword list, and dropped to 0.5 for Important sentences. The three lowest similarity rankings recorded values under 0.4 for plain text and keyword lists, with an additional decline to below 0.3 for important sentences.

Table 5: *BERT on Case 4.*

Ground Truth	Plain text		Keyword list		Important sentences	
	Nanda No	Similarity Score	Nanda No	Similarity Score	Nanda No	Similarity Score
039	**002**	*0.95*	**002**	*0.94*	**002**	*0.94*
002	**039**	*0.94*	**039**	*0.94*	**039**	*0.92*
	276	*0.90*	146	*0.89*	276	*0.86*
	047	*0.89*	276	*0.89*	108	*0.85*
	108	*0.89*	108	*0.89*	109	*0.85*
	•	•	•	•	•	•
	297	*0.78*	093	*0.79*	093	0.70
	129	*0.78*	007	*0.79*	007	0.70
	304	*0.77*	100	*0.79*	247	0.70

Table 6: *SBERT on Case 4.*

Ground Truth	Plain text		Keyword list		Important sentences	
	Nanda No	Similarity Score	Nanda No	Similarity Score	Nanda No	Similarity Score
039	**002**	*0.86*	**002**	*0.85*	**002**	*0.80*
002	**039**	*0.80*	**039**	*0.83*	**039**	0.73
	276	0.65	276	0.67	146	0.54
	108	0.62	108	0.67	276	0.52
	109	0.61	303	0.66	214	0.51
	•	•	•	•	•	•
	027	0.40	027	0.40	281	0.29
	011	0.39	064	0.40	083	0.29
	247	0.37	100	0.37	100	0.26

Proceeding to case 26, the results obtained using the BERT model are presented in Table 7, while those from the SBERT model are detailed in Table 8.

For case 26, the BERT experiment precisely identified four nursing issues corresponding to the ground truth, placing them within the top four rankings across the plain text, keyword list, and important sentences categories. Similarity scores for plain text and keyword list impressively exceeded 0.9 for all four nursing issues. Important sentences revealed similarity scores of 0.89 for two nursing issues. Remarkably,

Table 7: *BERT on Case 26.*

Ground Truth	Plain text		Keyword list		Important sentences	
	Nanda No	Similarity Score	Nanda No	Similarity Score	Nanda No	Similarity Score
108	**266**	*0.96*	**266**	*0.96*	**266**	*0.94*
132	**132**	*0.95*	**132**	*0.94*	**132**	*0.92*
266	**108**	*0.94*	**276**	*0.91*	**108**	*0.90*
276	**276**	*0.93*	**108**	*0.91*	**276**	*0.89*
	110	*0.93*	146	*0.89*	110	*0.88*
	.	.	.	.	.	.
	297	*0.80*	125	*0.79*	147	0.71
	129	*0.80*	299	*0.79*	093	0.71
	304	*0.79*	152	*0.78*	007	0.71

Table 8: *SBERT on Case 26.*

Ground Truth	Plain text		Keyword list		Important sentences	
	Nanda No	Similarity Score	Nanda No	Similarity Score	Nanda No	Similarity Score
108	**266**	*0.84*	**266**	*0.87*	**266**	*0.85*
132	**132**	*0.82*	**132**	*0.81*	**132**	*0.78*
266	**108**	*0.76*	**276**	0.71	**108**	0.67
276	**276**	*0.76*	**108**	0.71	**276**	0.67
	102	0.72	109	0.64	109	0.58
	.	.	.	.	.	.
	267	0.41	152	0.41	299	0.34
	299	0.38	027	0.40	125	0.34
	125	0.38	064	0.40	304	0.32

even the fifth-ranked issue, which was not a correct answer for this case, showed a similarity score above 0.8. Like case 4's BERT findings, the three lowest-ranked nursing issues maintained similarity scores above 0.7.

The SBERT analysis for, case 26, shown in Table 7, also correctly identified the four ground truth nursing issues within the top four ranks for plain text, keyword list, and Important sentences. While the predictions remained at the forefront, the results for plain text and keyword list indicated the top two issues had similarity scores above 0.8, and the subsequent two scored above 0.7. For important sentences, the highest-ranking issue achieved a similarity of 0.8, but the second rank fell to 0.7, with the third and fourth ranks at 0.6. Where the fifth rank did not correspond to the correct answer, similarity scores reduced to 0.7 for plain text, 0.6 for keyword list, and 0.5 for Important sentences. Echoing case 4's SBERT results, the three lowest ranks scored below 0.4 for plain text and keyword lists, and below 0.3 for important sentences.

The heatmap in Fig. 3 displays the results of the six tested patterns in rank order for case 26, using BERT and SBERT models. Yellow represents high similarity, while dark blue indicates low values.

In this study, we evaluated the effects of different pre-processing methods (plain text, keyword lists, and important sentences) and compared the performance of the BERT and SBERT models.

For BERT, no significant differences were observed among the various preprocessing methods. The model performed similarly regardless of whether plain text, keyword list, or important sentences were used.

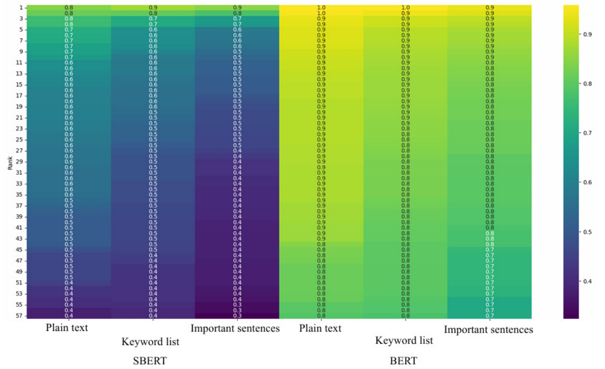


Fig.3: *Heatmap Comparison of SBERT and BERT Results for Case 26.*

However, BERT produced consistently high similarity scores for many nursing diagnoses, which made it challenging to distinguish and rank the most relevant ones. This widespread high similarity complicates determining the necessary diagnoses, as it does not clearly distinguish between relevant and irrelevant issues.

For SBERT, although the accuracy rate remained consistent across all preprocessing methods, the combination with important sentences showed the best similarity scores. The SBERT model exhibited higher similarity scores when matching the ground truth, while producing lower similarity scores for non-matching predictions. This transparent gradient from high to low similarity suggests that SBERT's predictions are more practical, as they more effectively eliminate irrelevant diagnoses and highlight the most relevant ones.

Overall, the results indicate that SBERT combined with important sentences provides the most precise and reliable outcome, particularly when considering the ability to differentiate between correct and incorrect diagnoses based on similarity scores. Conversely, BERT's widespread high similarity scores hinder the ability to rank diagnoses meaningfully.

6. DISCUSSION

6.1 Impact of Preprocessing Techniques and Models on Nursing Problem Prediction

The research presented herein validates the potential for predicting NANDA-I nursing diagnoses by correlating feature vectors, derived from Gordon's Functional Health Patterns, with nursing diagnoses. Utilizing a combination of subjective and objective data recorded by nurses, and employing an 11-dimensional feature vector X, our proposed method demonstrated considerable predictive accuracy. Adopting BERT and SBERT models, along with preprocessing techniques such as Plain text, keyword list, and important sentences, proved to be effective in accurately forecasting nursing issues across all evaluated cases.

SBERT accurately predicted ground truth diagnoses and differentiated from incorrect ones, as shown in Fig. 3. The observed discrepancy in performance between the BERT and SBERT models underscores the importance of domain-specific optimization in the nursing sector. Although BERT models are adept at capturing detailed contextual information within texts, they are not inherently designed to distinguish directly between the semantic similarities of different texts [26]. As a result, BERT produced uniformly high similarity scores for all nursing issues, demonstrating its strong contextual comprehension but revealing challenges in distinguishing specific nursing issues for prediction.

In contrast, the direct optimization for similarity calculations by SBERT enables a more nuanced capture of semantic nuances, yielding a more comprehensive array of similarity scores that more accurately reflect semantic distances between sentences. The SBERT model captured semantic nuances better than the BERT model, leading to more varied and accurate similarity scores. proves its effectiveness in discerning the relationships between nursing issues and their textual depictions, underlining the advantage of employing a specialized approach like SBERT for detailed analysis and predictions in niche areas such as nursing.

The preprocessing methodologies-plain text, keyword list, and important sentences—significantly influenced predictive accuracy across all patterns, underscoring the critical role of systematic text analysis in extracting meaningful patterns from nursing records. The application of cosine similarity, known for its efficacy in analysing short and domain-specific texts, showcased its utility in deepening the understanding of nursing records. This demonstrates the potential of cosine similarity to provide deeper insights into nursing documentation.

Nonetheless, this investigation was constrained by the modest dataset size of 29 cases, highlighting the challenges associated with leveraging actual hospital data due to privacy considerations. The limited dataset size posed challenges, but the models could still provide accurate diagnostic predictions. Due to the small data size, the study could not produce comprehensive predictions for all NANDA-I diagnoses, limiting the scope of the analysis. Nursing problems that appeared in only one case could not be effectively learned and predicted, resulting in the model being able to accurately predict only those diagnoses that appeared multiple times in the dataset. This data imbalance may also have led to overfitting in specific diagnostic categories, limiting the model's generalization capabilities. Addressing these challenges will require more extensive, more diverse datasets in future research. The data imbalance—from nursing diagnoses appearing in just one case to those in ten—suggests variability in the how each nursing diagnosis label's feature vector reflects individual or multiple cases. The data used were from simulated cases, which likely contained minimal noise, contributing to the accuracy of the model's predictions. This indicates that feature vector generation via plain text may have been particularly effective owing to the limited data scope. Conversely, important sentences yielded lower similarity scores for test cases than plain text, yet demonstrated decreased similarity for incorrect nursing issues, emphasizing the precision of BERT and SBERT models in excluding irrelevant nursing issues. Generating feature vectors using plain text was adequate, given the limited data range. At the same time, meaningful sentences reduced similarity scores for incorrect nursing problems, highlighting the models' accuracy in excluding irrelevant issues.

Given the nature of nursing issues, overprediction could lead to inefficiencies in nursing workload, whereas underprediction might degrade the quality of care. Therefore, offering a balanced array of predictive candidates is imperative for optimizing the nursing workload and enhancing the quality of patient care.

A balanced set of candidate predictions is crucial to optimize workload and improve patient care. Future studies should experiment with larger datasets that include diverse nursing records.

6.2 Utilizing Data Science Techniques in the Nursing

The results of this study suggest that effectively leveraging the information documented in daily nursing records by nurses can lead to more accurate predictions of nursing diagnoses. This suggests that the nursing process, from assessment to planning, could be made more efficient and accurate with AI support. Utilizing NLP and AI technologies for selecting nursing diagnoses and planning nursing care could support decision-making, thus alleviating the burden on individual nurses faced with the challenge of nursing shortages and enhancing the quality of patient care.

Future research should focus on constructing a larger dataset incorporating diverse nursing records and further optimizing and evaluating NLP models tailored for the nursing domain. Moreover, automating the manual tasks of categorizing data patterns and assigning NANDA-I nursing diagnoses is necessary for system development.

Conducting empirical studies to verify the applicability and effectiveness of these models in actual nursing settings is crucial. Through such efforts, this study aims to utilize AI and NLP technologies effectively as decision-support tools in nursing, ultimately improving the quality of patient care.

7. CONCLUSION

This study has substantiated the feasibility of employing advanced NLP models, specifically BERT and SBERT, for predicting NANDA-I nursing diagnoses. By integrating feature vectors derived from MGFHP with textual data meticulously recorded by nurses, we have demonstrated enhanced accuracy in diagnosing nursing issues across multiple test cases. This underscores the critical importance of domain-specific adaptations in model training and the rigorous analysis of textual data within the nursing domain. Our results highlight the transformative potential of NLP technologies in refining nursing diagnostics and improving care planning processes. The study advocates for expanding data diversity and further model refinement engineered explicitly for applications within nursing to bolster predictive performance. Moreover, the automation of data classification and the assignment of diagnoses are proposed as innovative approaches to streamline nursing operations and enhance the quality of patient care. Future research should prioritize the empirical validation of these models within real-world nursing settings to confirm their utility as practical decision-support tools.

Future investigations should also focus on developing AI frameworks capable of autonomously identifying urgent cases through the analysis of real-time patient data and integrating predictive models to assist nursing professionals in their daily duties. Such technological advancements could significantly empower nursing personnel, enabling more informed and strategic decision-making processes, and ultimately raising the standard of nursing care. Future research should focus on increasing the number of simulated cases to incorporate more diverse data. Specifically, multiple nurses could be asked to apply nursing diagnoses to these simulated cases, generating a richer and more varied dataset for model training. This step will help improve the model's generalization capabilities. Eventually, using accurate clinical data will be essential for validating the model's effectiveness in actual clinical environments. By following these steps, the model can be adapted for practical use in clinical settings, improving both the robustness and utility of nursing diagnostics, and potentially enhancing the accuracy and efficiency of nursing care.

AUTHOR CONTRIBUTIONS

Investigation of methods, Takahiro Kubo, Virach Sornlertlamvanich and Thatsanee Charoenporn; Data collection, Takahiro Kubo; Experiment, Takahiro Kubo and Virach Sornlertlamvanich; Discussion, Takahiro Kubo, Virach Sornlertlamvanich and Thatsanee Charoenporn; writing—original draft preparation, Takahiro Kubo; writing—review and editing, Takahiro Kubo, Virach Sornlertlamvanich and Thatsanee Charoenporn; visualization, Takahiro Kubo; supervision, Virach Sornlertlamvanich.

All authors have read and agreed to the published version of the manuscript.

References

[1] State of the World's Nursing 2020: investing in education, jobs and leadership, Genève, Switzerland: World Health Organization, 2020.

[2] Study Group on the Supply and Demand of Health Care Workers, Subcommittee on the Supply and Demand of Nursing Staff, Interim Summary, 2019 (in Japanese)

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

[4] S. Mitha *et al.*, "Natural Language Processing of Nursing Notes: A Systematic Review," *Stud Health Technol Inform*, pp.62–64, 2021.

[5] Herdman, T. Heather, S. Kamitsuru and C. T. Lopes, *Nanda International Nursing Diagnoses*, 12th ed., Thieme Medical Pub, 2021.

[6] Japan Nursing Association Public Interest Incorporated Association. Revised edition: Explanation of key terms related to nursing. Japan Nursing Association Public Interest Incorporated Association, p.15, 2023.(in Japanese)

[7] The 9th and 10th terms of the Japanese Society of Nursing Science's Committee for the Examination of Nursing Terminology. A glossary of important terms that make up the field of nursing, p.7, 2011. (in Japanese)

[8] American Nurses Association, *Nursing: Scope and Standards of Practice*, 4th ed., American Nurses Association, 2021.

[9] C. M. Wagner, H. K. Butcher and M. F. Clarke, *Nursing Interventions Classification (NIC)*, 8th ed., Elsevier, 2023.

[10] S. Moorhead, E. Swanson and M. Johnson, *Nursing Outcomes Classification (NOC) Measurement of Health Outcomes*, 7th ed., Elsevier, 2023.

[11] D. Shibata, Z. Y. Kawa, K. Shimamoto, E. Shinohara and E. Aramaki, "Pain Expressions Extraction with Pre-training BERT Based on Clinical Note," *Medical Informatics*, vol. 40, no. 2, pp.73–82, 2020. (in Japanese)

[12] H. Abe, S. Hirano and S. Tsumoto, "Evaluation Temporal Models Based on Association Mining from Medical Documents," *Japan Journal of Medical Informatics*, vol. 27, no. 1, pp33–38, 2007. (in Japanese)

[13] E. Matsumura, Y. Shimizu, S. Tachibana and E. Ueno, "The Effect of PNS Developed University of Fukui Hospital -Ward Nurses's Recognition After One Year to the Innovation of PNS," *Research Journal of the University of Fukui School*

of Medicine, vol. 16, no. 1, pp. 37–46, 2016. (in Japanese)

[14] S. Mitha *et al.*, "Natural Language Processing of Nursing Notes: An Integrative Review," *CIN: Computers, Informatics, Nursing*, vol. 41, no. 6, pp.1–8, 2023.

[15] M. J. Press *et al.*, "Postdischarge Communication Between Home Health Nurses and Physicians: Measurement, Quality, and Outcomes," *J Am Geriatr Soc*, vol. 63, no. 7, pp.1299–1305, 2015.

[16] K. Woo, *et al.*, "Exploring prevalence of wound infections and related patient characteristics in homecare using natural language processing," *Int Wound J*, vol. 19, no. 1, pp. 211–221, 2022.

[17] P. Hamet and J. Tremblay, "Artificial intelligence in medicine," *Metabolism*, vol. 69, pp.36–40, 2017.

[18] J. Shi *et al.*, "Global output on artificial intelligence in the field of nursing: A bibliometric analysis and science mapping," *J Nurs Scholarship*, vol. 55, no. 4, pp.853-863, 2022.

[19] E. J. Topol, "High-performance medicine: the convergence of human and artificial intelligence," *Nature Medicine*, vol. 25, no. 1, pp.44–56, 2019.

[20] A. Comelli, L. Agnello and S. Vitabile, "An ontology-based retrieval system for mammographic reports," *2015 IEEE Symposium on Computers and Communication (ISCC)*, Larnaca, Cyprus, pp. 1001-1006, 2015.

[21] P. Sitikhu, K. Pahi, P. Thapa and S. Shakya, "A Comparison of Semantic Similarity Methods for Maximum Human Interpretability," in *2019 Artificial Intelligence for Transforming Business and Society (AITB)*, pp. 1–4, 2019.

[22] Lab TN, " Pretrained Japanese BERT models released," Accessed: May. 24, 2022, [Online.] Available from: `https://www.nlp.ecei.tohoku.ac.jp/newsrelease/3284/`

[23] N. Reimers and I. Gurevych, "Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks," [Online]. arXiv [cs.CL]. 2019. Available from: `http://arxiv.org/abs/1908.10084`

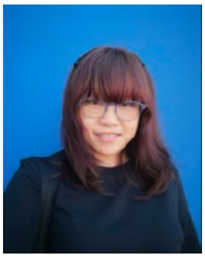
[24] Y. Kawazoe, D. Shibata, E. Shinohara, E. Aramaki and K. Ohe, "A clinical specific BERT developed using a huge Japanese clinical text corpus," *PLoS One*, vol. 16, no. 11:e0259763, 2021.

[25] C. Sun, X. Qiu, Y. Xu and X. Huang, "How to Fine-Tune BERT for Text Classification?," *Chinese Computational Linguistics*, Springer International Publishing, pp. 194–206, 2019.

[26] T. Zhang, V. Kishore, F. Wu, K. Q. Weinberger and Y. Artzi, "BERTScore: Evaluating Text Generation with BERT," *arXiv*, arXiv: 1904.09675. 2019.

Kubo Takahiro is an Assistant Professor at the Faculty of Nursing, Musashino University. He holds a bachelor's and a master's degree in nursing and is currently pursuing a doctoral degree in data science. After becoming a licensed registered nurse and gaining clinical experience in a hospital setting, he transitioned to academia. His research focuses on integrating nursing with data science and exploring healthcare technology.

Virach Sornlertlamvanich received the B.Eng. and M.Eng. degrees in mechanical engineering from Kyoto University, Kyoto, Japan, in 1984 and 1986, and the Ph.D. degree in computer engineering from the Tokyo Institute of Technology, Tokyo, Japan, in 1998. He was a recipient of the "National Distinguished Researcher Award" in Information Technology and Communication from the National Research Council of Thailand, in 2003, followed by the "ASEAN Outstanding Engineering Achievement Award" from the ASEAN Federation of Engineering Organizations (AFEO), in 2011, and was esteemed as "The Researcher of the Year 2001" by the Nation Newspaper, in 2001. He was also recognized with an ''Outstanding Alumni Award" by Tokyo Tech Alumni Association (Thailand Chapter), in 2021.



Thatsanee Charoenporn is a Professor in the Faculty of Data Science at the Asia AI Institute, Musashino University. She holds a B.A. and M.A. in Linguistics and a Ph.D. in Information Technology. She has previously held research and project leadership roles at the National Institute of Information and Communications Technology (Japan), the National Electronics and Computer Technology Center (Thailand), and Burapha University (Thailand). Prof. Charoenporn has extensive experience in multidisciplinary research and projects. Her current research focuses on language resources, cross-cultural communication, smart education, healthcare technology, and social innovation.