

Article

Not peer-reviewed version

Methodology of Labeling According to 9 criteria of DSM-5

Geon-Ju Lee , Da-Bin Park , [Ha-young Oh](#) *

Posted Date: 12 June 2023

doi: 10.20944/preprints202306.0756.v1

Keywords: DSM-5; Natural Language Processing; Morphological analysis; labeling; classification; Embedding; depressive disorders



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Methodology of Labeling according to 9 Criteria of DSM-5

Geon-Ju Lee ¹, Da-Bin Park ² and Ha-Young Oh ^{3,*}

¹ Department of Mathematics, Sungkyunkwan University, Seoul, 03063 Korea; dlrjsw179@naver.com

² Applied Artificial Intelligence, Sungkyunkwan University, Seoul, 03063 Korea; dabin3178@gmail.com

³ College of Computing and Informatics, Sungkyunkwan University, Seoul, 03063 Korea

* Correspondence: hyoh79@gmail.com; Tel.: +82-2-583-8585

Abstract: Depressive disorder is a disease that causes a decrease in daily function, and the prevalence of depressive disorder in Korea is the highest among OECD countries. However, due to lack of manpower and negative social awareness, timely treatment is not being performed properly. As a solution to this, AI counseling chatbots are emerging, and it is essential to use labeled datasets to create chatbots. Currently, many countries around the world, including Korea, use the Diagnostic and Statistical Manual of Mental Disorders Fifth Edition, DSM-5, but there are no nine labeled datasets according to the DSM-5 depression disorder diagnosis criteria. This study collects sympathetic datasets, analyzes morphemes using Kind Korean Morpheme Analyzer and augments and builds a word dictionary using Word2Vec. As a result, we have labeled the dataset with nine criteria for diagnosing DSM-5 depressive disorder, which will enhance the performance of the counseling chatbot.

Keywords: DSM-5; natural language processing; morphological analysis; labeling; classification; Embedding; depressive disorders

1. Introduction

Depressive disorder is a disease that causes various cognitive and mental and physical symptoms, resulting in a decrease in daily function. Severe depressive disorder can lead to suicidal thoughts as well as repetitive thoughts of death [1]. The prevalence of depressive disorders in Korea is 36.8%, and 4 of every 10 people in Korea suffer from depressive disorders or depression, which is the highest among OECD countries [2].

The fifth edition of the Diagnostic and Statistical Manual of Mental Disorders Fifth Edition (DSM-5) is the official reference of the American Psychiatric Association (APA). The DSM-5 is also used around the world, and depressive disorders are included in its classification of mental disorders. Depressive disorders are determined by comprehensively considering the nine diagnostic criteria of the DSM-5, psychiatrists' clinical experiences, clinical interviews, psychological test results, and treatment progress.

Fear caused by negative social perceptions of mental disorders can prevent counselors, including psychiatrists, from being treated for depressive disorders. In addition, it can be difficult to intervene in a timely manner due to the lack of infrastructure to cope with depressive disorders when they are accompanied by a physical disease [3]. Accordingly, natural language processing-based artificial intelligence (AI) services that can be consulted in a timely manner are emerging [4,5]. Among them, conversational AI agents known as chatbots can help predict mental illness conditions by consulting a counselor's textual records.

The term chatbot refers to a robot that can converse with users. Natural-language processing and machine learning are essential for chatbots to carry on meaningful and complex conversations. In this case, supervised machine learning is the primary tool, and it is essential to label datasets to ensure correct answer are used during training sessions. In addition, because perceived empathy with users can lead to a positive evaluation of a chatbot's responses, it is effective to use empathy datasets [6].

This paper labels an empathetic dataset according to DSM-5 depressive disorder diagnostic criteria. A wellness consultation dataset provided by AI Hub [7] and a dataset obtained by crawling the depression subreddit channel of the Reddit internet forum were used. In addition, the corresponding datasets were classified and labeled through the use of a word dictionary that classifies words expressing the elements of depressive disorders according to nine criteria.

Currently, none of the datasets on the market are labeled according to the DSM-5 depressive disorder diagnostic criteria. Accordingly, this paper labeled the dataset using DSM-5 depression disorder diagnostic criteria, which are widely used by the medical community in various fields. In medicine, the accuracy of diagnosis can be improved through the use of corresponding datasets and machine-learning models, and in natural-language processing, the criteria can be used in chatbots programmed for depression counseling. For chatbots specifically, empathy datasets and actual counseling datasets can help elicit positive evaluations by revealing an empathetic attitude toward the user.

2. Previous Research

Kang Seung-sik et al. [8] analyzed the vocabulary usage characteristics of the short messaging service (SMS) character corpus, which is a spoken corpus, the Naver movie review corpus, and the Korean written primitive corpus, which is a written corpus. To measure the strength of the performance, a discriminatory performance vocabulary analysis methodology was used, and adjectives that appear mainly in the spoken and written forms were identified. In addition, a word-embedding technique was used to automatically build an emotional vocabulary dictionary based on adjectives with high performance intensity. A total of 343,603 emotional vocabulary dictionaries were automatically constructed using this technique. However, related research has not been able to manually select a large number of automatically constructed professional vocabularies.

Kim et al. [9] proposed a machine learning-based emotional-analysis system that detects user depression through SMS messages. Using an emotion dictionary of words related to depression and other emotions, search keywords were selected from text data on Twitter to build a learning dataset consisting of 1,297 sentences that express depression-related emotions and 1,032 sentences without them. Finally, a circulatory neural network, short- and long-term memory, and gate circulation units were compared and evaluated. Among them, a model based on gate circulation units achieved an accuracy of 92.2%. However, Kim et al. limited depression-related emotions to "sadness."

Seo et al. [10] collected conversation datasets from 2016 to February 2019 from users who tended to be depressed and those who were not, and analyzed the subject and vocabulary characteristics of the dataset. The periods before and after the depression trend were compared, and the depression trend was analyzed by period. In addition, topic modeling, simultaneous word analysis, and emotional analysis were used to identify the characteristics and differences among topics of conversation of users with respect to depression and non-depression trends, and to analyze the vocabulary used. However, the Seo et al. team's biased thoughts may have affected the outcome when classifying words associated with depression and non-depression.

Chin et al. [11] collected conversation datasets related to depressed emotions in chats between people and chatbots. They analyzed their data using text-mining techniques to identify conversation topics related to depression-related emotions. In addition, through qualitative analysis, the types of depression that users described when confiding in chatbots were categorized and classified, and differences were identified by comparing them with depression-related data on Twitter.

3. Related Techniques

This section may be divided by subheadings. It should provide a concise and precise description of the experimental results, their interpretation, as well as the experimental conclusions that can be drawn.

3.1. Kind Korean Morpheme Analyzer

Morphological analysis refers to classifying sentences into morphemes and grasping the linguistic structure, and can distinguish roots, prefixes, suffixes, and parts of speech. The little morpheme analyzer [12] is a Korean morpheme analyzer that is less sensitive to spacing errors. For speed, an adjacent condition-inspection method using a mood dictionary was used, and for quality, a probability model based on the Heuristic and Hidden Markov Model was used. In the case of a heuristic model, morphemes were analyzed by considering the characteristics of the language, such as the part-verb of the previous word and the part-verb of the next word. Morpheme analysis results were then sorted to ensure the most appropriate information was obtained. A hidden Markov model is a probability model with both hidden and observed states. In the little morpheme analyzer, each morpheme is modeled as a hidden state and the parts of the morpheme are modeled as observable. Morpheme analysis estimates the probability of the appearance of morphemes. Figure 1 depicts an example of the morpheme analysis process using a little morpheme analyzer.

```
['아버지', '가', '방', '에', '들어가', '입니다', '.']
(['아버지', 'NNG'),
 ('가', 'JKS'),
 ('방', 'NNG'),
 ('에', 'JKM'),
 ('들어가', 'VV'),
 ('입니다', 'EFN'),
 ('.', 'SF')]
```

Figure 1. Example of the Kind Korean Morpheme Analyzer.

3.2. Word2Vec

Word2Vec [13] is a word-embedding technology presented by Google in 2013. Word2Vec distributes the meaning of a word across multiple dimensions, expresses it as a vector, and calculates the similarity between word vectors. At this time, analysis of words that appear in similar contexts assumes that they are similar. In addition, after making the word blank, the researchers guessed the word to place in the blank using the word adjacent to the blank in the sentence and one of two methods. Continuous bag of words (CBOW) methods leave the surrounding words and middle words blank, and Skip-Gram methods predict surrounding blank words as intermediate words. An example of the Word2Vec expression is shown in Figure 2.

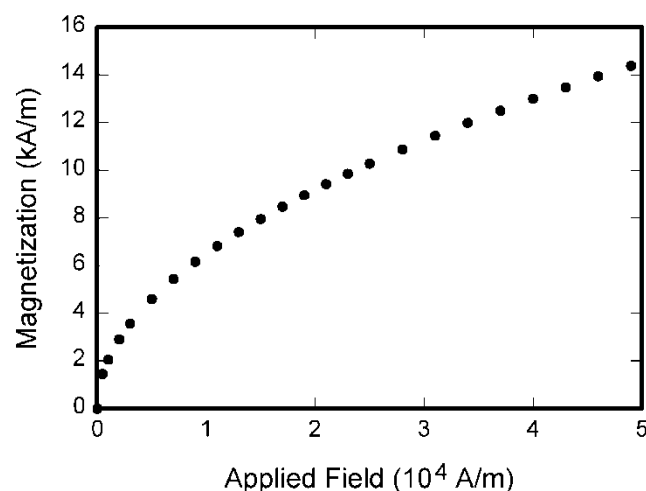


Figure 2. Magnetization as a function of applied field.

3.3. Dataset Labeling and Classification

Dataset labeling provides information about the data, and dataset classification assigns data to groups. In general, machine-learning models use labeled datasets to solve classification problems. For machine-learning models to learn efficiently and perform predictions, high-quality labeled datasets are required. Datasets can be labeled manually or automatically. Manual labeling assigns datasets to experts and other categories and is often highly accurate, but can be costly, both in terms of money and time. Semi-graphic learning combines hand-labeled and non-labeled data, which can reduce labeling costs at the expense of mislabeled or missing data.

4. Dataset

4.1. Dataset Description

“Wellness counseling data” were created by processing records of mental health counseling at Severance Hospital in Sinchon. Only data related to mental symptoms were used. A total of 5,231 questions and 1,034 responses were included. The “intent” column has 19 labels: depression, sadness, loneliness, anger, lethargy, emotional control abnormalities, loss, loss of appetite, appetite, insomnia, anxiety, fatigue, guilt, concentration, self-confidence, despair, suicide impulse, and anxiety.

In the wellness counseling data, the average length of a counselor’s utterance is 28 characters (minimum 2, maximum 117), and the average length of a counselor’s utterance is 31 characters (minimum 5, maximum 67). Figure 3 shows the results of visualizing words that mainly appear in each counselor’s speeches and responses as a word cloud.



Figure 3. Wordcloud of ‘utterance (2 차)’ column and ‘response (공감)’ column.

Data from the r/depression subreddit channel were extracted from Reddit, a mega-social community site in the United States. In this paper, 30,000 posts and comment data from 2010 to 2016 were randomly extracted. In the “depression” subreddit, any violation of the notice and detailed rules will be deleted regardless of the number of recommendations. The notice and detailed rules are shown in Figure 4.

Community Basics	r/depression Rules
<p>We offer a peer support space for anyone dealing with a depressive disorder in themselves or someone close to them.</p> <p>Please stay on topic. Depression is both important and difficult to talk about so focus is essential. Posts here need to be support requests specifically related to depression, and comments need to be supportive of the OP.</p> <p>If you want to talk about thoughts or risk of suicide, please post at r/SuicideWatch; if you've lost someone to suicide, r/SuicideBereavement is the best community to get support. It might seem that we have a lot of rules, but we've found they're all necessary to maintain as much emotional and physical safety as possible. Most people are surprised by at least some of our policies so please read all of them carefully before jumping in.</p> <p>Please click "report" to let us know of any inappropriate content you see here - we'd like to know and handle it as soon as we can.</p> <p>If your post or comment is not appearing, it may have been removed for a rule violation or it may be stuck in the filter. Please message us and we'll look into it.</p> <p>We are not a crisis service. We can't guarantee an immediate response, and there are times when this subreddit is relatively quiet. This does not mean no one cares. If you need to talk to someone at once, you may want to take a look at the hotlines list from r/SuicideWatch</p> <p>Note: the Android Reddit app currently has a bug that causes wikis not to work - if links do not open in the app, use your browser to open the links instead</p>	<ol style="list-style-type: none"> 1. Posts must request personal support for depressive disorders in yourself, or someone close to you 2. All replies must be empathetic and responsive to OP. No tough love or debate. 3. Offer help ONLY via public comments. No "I'm here to help" posts or PM/chat requests 4. No low-content posts or responses 5. Don't role-model or put yourself above others. No achievement/advice posts or AMAs 6. No general uplifting content of any kind, including "it gets better" or other unkeepable promises 7. Do not give or request clinical advice or advocate for or against treatments & self-help strategies 8. NO methods for any type of self-harm, or validating self-destructive intent 9. Don't encourage rule-breaking or off-topic content 10. No activism, debate or unapproved surveys including casual or general questions and polls 11. Absolutely no self-promotion of any kind, including creative writing 12. Do not request or offer money, goods, or services

Figure 4. Community basics and rules of r/depression.

Users of r/depression can only ask for help with depressive disorders for themselves or someone close to them, and all answers should be sympathetic. In addition, answers to posts asking for help are only allowed in public comments, and personal answers and posts are not allowed. Responses without meaningful content are not allowed, and responders should not pose as a role model or someone with a superior perspective. Meaningless encouragement, promises that cannot be kept, and medical knowledge, advocacy, and opposition to treatments are also not permitted. In addition, it is impossible to certify self-harm methods and self-destructive content, and content that deviates from the subject or goes against the rules should not be encouraged. Finally, it is impossible to request and provide discussions, unauthorized surveys, self-promotion, and money, goods, services, etc.

4.2. Dataset Preprocessing

In the case of wellness counseling data, the “utterance” column, which involves counselors’ utterances, and the “response” column, which involves counselors’ responses, were used. After that, the missing value was deleted. In the case of r/depression data, sentences containing adjectives on either side of a “/”, such as “depressed/anxious,” were deleted. An example of a sentence is shown in Figure 5.

depressed/anxious for years and years
I always have been on the depressed/anxious side, and very hard on myself, especially about my looks
what are your experiences with being or having a depressed/anxious parent?
when i am with her i feel more relaxed and less depressed/anxious
maybe i should just focus on getting a simply job so that i can do it even if depressed/anxious
I am depressed/anxious and it is just wearing me down
additionally, depressed/anxious people tend to be sensitive to things that may seem insignificant to others so for as long as I can remember i have been having depressed/anxious/suicidal feelings almost all the time
I get depressed/anxious on sundays/mondays really bad
sometimes i feel depressed/anxious and can't stop thinking about the future and what i have to do and what i can't do

Figure 5. r/depression data examples including “depressed/anxious.”

Translations were made using the Google Translate API [14]. An example of a translation result is shown in Figure 6.

나는 중독성 성격이 없어서 지난 2년 동안 항우울제에 트라마돌을 보충하고 있습니다.
나는 가끔 술에 취해 보인다
하지만 지난 28개월 동안 나는 높은 스트레스를 유지했고 일을 빠르게 진행했으며 그 일에 탁월했습니다
나는 운전이 문제가 있었습니다. 그것은 아마도 내 모든 에너지를 소모하는 일과 더 관련이 있을 것입니다.
그것은 삶을 변화시켰습니다.
tramadol은 몇 시간 후에 작동하고 약 12 지속됩니다.
당신은 그것을 온라인으로 얻을 수 있습니다
아편류와 유사하지만 부작용이 훨씬 적고 머리가 흐릿하거나 입이 마르는 등의 현상이 없습니다.
중독성이 있다는게 큰 단점
너무 유용하고 잘 작동하고 빠르게 작동하는 것이 중독성이 있다는 것은 정말 개년이라고 생각합니다.

Figure 6. Translation examples of r/depression data.

5. Dictionary of Expression Words

A word dictionary was created based on the nine DSM-5 diagnostic criteria for depressive disorder. Medical staff working in the Department of Mental Health Medicine collected expressions for each of the nine depressive disorder criteria and generated a dictionary, as shown in Figure 7.

1. 하루의 대부분, 거의 매일 지속되는 **우울한 기분**
2. 하루의 대부분, 거의 매일 거의 모든 일상 활동에 대한 흥미나 즐거움이 현저히 감소
3. 식이 조절 없이 **체중 감소 또는 증가**, 거의 매일 **식욕 감소 또는 증가**
4. 거의 매일 **불면 또는 과수면**
5. 거의 매일 정신 흥분 또는 지체
6. 거의 매일 **피로 또는 에너지 상실**
7. 거의 매일 단순한 자기 비난, **무가치감** 또는 과도하고 부적절한 **죄책감**
8. 거의 매일 **사고력과 집중력의 감소**, 결정 곤란
9. 죽음에 대한 반복적인 생각, 계획 없는 반복적인 **자살 생각** 또는 시도나 자살을 시도하려는 상세한 계획

Figure 7. DSM-5 diagnosis criteria for depressive disorder.

The number of words in the word dictionary for each criterion is 59, 45, 44, 33, 37, 27, 45, 26, and 47. Figure 8 provides an example of the first of the nine criteria: “most of the day, almost daily depressed mood.”

1. 하루 중 대부분 그리고 거의 매일 지속되는 우울 기분에 대해 주관적으로 보고 (예, 슬픔, 공허하다 또는 절망감) 하거나 객관적으로 관찰됨(예, 눈물흘림)
우울한 기분 (슬픔, 절망감, 무력감, 무가치감)
가망이 없다
가슴 아프다
가치 없다

Figure 8. Example of the word dictionary of the first DSM-5 diagnosis criteria for depressive disorder.

6. Methodology

6.1. Morphological Analysis

Each sentence of the preprocessed data was analyzed using a little morpheme analyzer. The preprocessed word dictionary was also morphemically analyzed and stored for each of the nine criteria. An example of each result is shown in Figure 9.

```
[('아이', 'NNG'),
 ('가지', 'YY'),
 ('고', 'ECE'),
 ('나', 'NP'),
 ('서', 'JKM'),
 ('우울하', 'YY'),
 ('어', 'ECS')]
[('덧', 'XPN'), ('없다', 'UN')]
```

Figure 9. Results of morphological analysis of datasets and word dictionaries.

An example of the results of morphological analysis of the entire word dictionary is shown in Figure 10.

1. : [['우울', '기분', '슬프', '절망감', '우려감', '우가', '치감']
2. : [['가라앉', '['감정', '안', '느껴지'], ['감흥', '없']
3. : [['끓', '['꾸역꾸역'], ['끼니', '거르'], ['다이어트'], ['더부룩']
4. : [['수면', '초기', '중기', '말기'], ['깊', '뭇', '자']
5. : [['사람', '눈치', '채', '정도', '평소', '말', '행동', '느리', '지']
6. : [['전반적', '이', '신체증'], ['갈리', ['고되'], ['근육통']
7. : [['나', '때문', '이'], ['나', '자신', '실망'], ['나', '가망', '없']
8. : [['결정', '장애'], ['결정', '힘들'], ['계속', '힘들']
9. : [['자살', '또는', '자해'], ['고통', '벗어나', '싶']

Figure 10. Results of morphological analysis of word dictionaries.

6.2. Morphological Screening

The results of the DSM-5 morpheme analysis were applied to the part-time tag table of the little morpheme analyzer (Figure 11). To extract only words with specific meanings, only parts, general adverbs, and body-language prefixes corresponding to body language, proverbs, and roots were selected.

대분류	새종 품사 태그		심경첩 품사 태그		KKMA 단일 태그 V 1.0		확률태그	저장시점			
	태그	설명	Class	설명	뜻음 1	뜻음 2					
체언	NNG	일반 명사	NN	명사	N	NNG	보통 명사	NNA	noun.dic		
	NNP	고유 명사	NX	의존 명사		NNP	고유 명사				
	NNB	외존 명사	UM	단위 명사		NNB	일반 의존 명사			NNS	
	NR	수사	NU	수사		NNM	단위 의존 명사			NR	simple.dic
	NP	대명사	NP	대명사		NR	수사			NR	
	용언	VV	동사	VV		동사	V			VV	동사
VA		형용사	AJ	형용사	VA	형용사					
VX		보조 용언	VX	보조 동사	VXV	보조 동사		VX			
VCP		공정 지칭사	CP	서술격 조사 '이다'	VXA	보조 형용사		VCP	taw.dic		
VCN		부정 지칭사			VCP	공정 지칭사, 서술격 조사 '이다'		VCN			
					VCN	부정 지칭사, 형용사 '아니다'		VCN			
관형사	MM	관형사	DT	일반 관형사	M	MD	일반 관형사	MD	simple.dic		
			DN	수 관형사		MDN	수 관형사				
부사	MAG	일반 부사	AD	부사	MA	MAG	일반 부사	MAG	simple.dic		
	MAJ	접속 부사				MAC	접속 부사			MAC	
김단사	IC	감탄사	EX	감탄사	I	IC	감탄사	IC			

Figure 11. Tag table of KKMA.

6.3. Data Classification

Each sentence of the consultation dataset was searched to determine if the words were included in the dictionary, and automatic classification was performed accordingly in a cloud-based Jupiter laptop environment using a Tesla T4 graphical processing unit accelerator. An example of the classification process is shown in Figure 12.

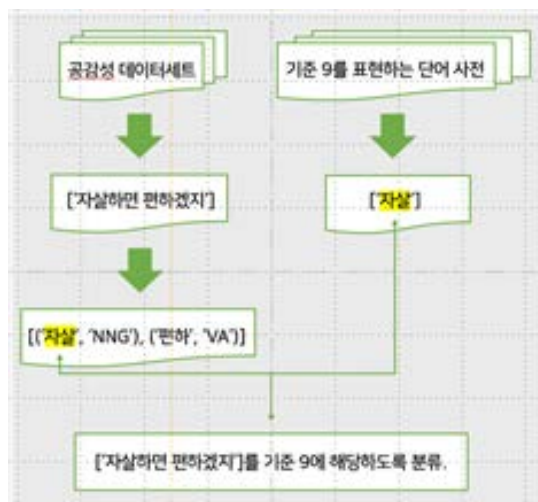


Figure 12. Example of a classification process.

6.4. Data Labeling

Classified data sets were labeled according to the DSM-5 depression disorder diagnostic criteria. For example, if the word dictionary for the third diagnostic criterion, “loss or gain weight without diet, loss or increase of appetite almost every day,” contained a “pig” in the counseling dataset sentence, the sentence was given a label of number 3. Figure 13 shows the results of labeling wellness consultation data according to the nine diagnostic criteria and visualizing the frequency.

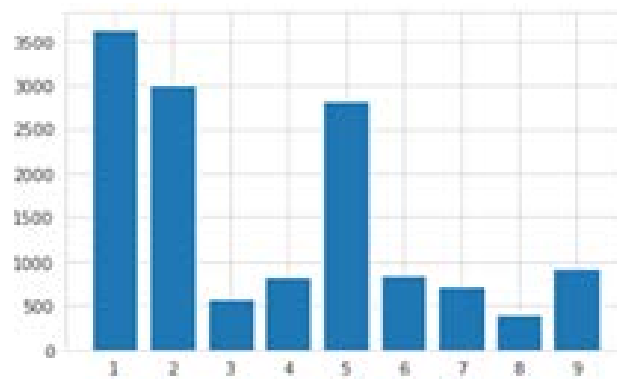


Figure 13. Visualization of dataset classification.

6.5. Data Evaluation

In 5.4, it was confirmed that the number of datasets belonging to diagnostic criteria 3, 4, 6, 7, and 8 was small. The similarity of words for the corresponding diagnostic criteria was therefore searched using Word2Vec. In the case of Word2Vec’s hyperparameter, the window was set to 5, and the five words to the left and right of the word were predicted using the context. In addition, the min_count parameter specified the minimum frequency to be used for learning, and words that appeared fewer than 5 times were excluded from learning. Finally, the CBOW algorithm was used by assigning a value of 0 to sg, which is a parameter for designating the algorithm. Figure 14 shows the results of searching for words similar to “lonely,” “terrible,” and other terms.



Figure 14. Words similar to “lonely” and “terrible.”.

The words and results of visualizations similar to “tough” are shown in Figure 15.

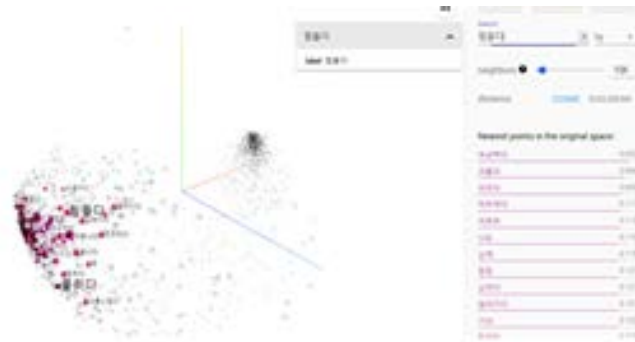


Figure 15. Visualization and words similar to “tough.”.

Similar words printed in this way were secondarily selected, and the same process was performed by adding them to the DSM-5 expression word dictionary.

7. Result

7.1. Comparing the Frequency of Each of the 9 Criteria

Figure 16 compares the frequency of wellness counseling data before and after word pre-augmentation, and Figure 17 compares the frequency of r/depression data from Reddit before and after augmentation.



Figure 16. Comparison of the wellness dataset classification.



Figure 17. Comparison of the Reddit dataset classification.

Combining the wellness consultation data and Reddit data produced the following frequency visualization.

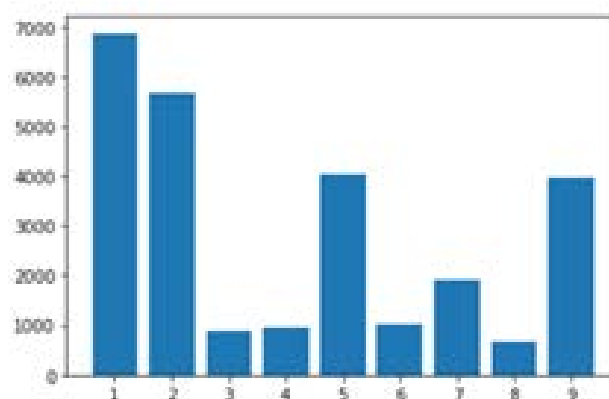


Figure 18. Comparison of entire dataset classification.

7.2. Frequency Result Inference

No significant difference was evident in the visualization results before and after augmentation of the word dictionary because many words were not added during dictionary enhancement. In addition, there was no significant difference in the counseling dataset collected in this study due to the lack of data corresponding to the label.

However, criteria 1 and 2 are “most of the day, almost every day of depression,” and “most of the day, almost every day of interest or pleasure in almost all daily activities is significantly reduced.” This indicates that the expression words are more ambiguous than other criteria and are often used. Compared with other criteria, the number of expressions such as “loss or gain weight without diet control, loss or increase of appetite almost every day” is small, but clear. However, diagnostic criteria 1 include “tears,” “sad,” and “disappointed,” which are ambiguous and frequently used in other contexts. As a result, we concluded that the number of data is inevitably high when the expression word is an ambiguous and widely used label.

8. Conclusions

This paper proposes a methodology for classifying and labeling datasets based on DSM-5 diagnostic criteria for depressive disorder. Wellness counseling data and Reddit crawling data consisting of empathy-related content were used to augment AI services for depressive counseling. A word dictionary built on medical knowledge was used and we analyzed the dataset and word dictionary using a little morpheme analyzer, with the resulting dataset labeled 1 to 9. In addition, if the number of words was insufficient, the word dictionary was augmented using Word2Vec.

A word dictionary built by medical staff and based on the DSM-5 depression disorder diagnosis criteria used by the APA was employed in this study. The proposed methodology can therefore be used in both medical care and natural language processing. By using datasets classified by proven diagnostic criteria, useful answers can be expected from natural language processing-based AI services such as chatbots. In addition, this approach can lead to a combination of AI and medical services that may be able to help solve problems related to the lack of counselors and difficulty supplying timely intervention.

Author Contributions: Conceptualization, methodology, experiment, visualization, formal analysis, G.J.L.; writing and editing, D.B.P.; supervision, H.Y.O. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Research Foundation of Korea grant funded by the Korea government (No. NRF-2022R1F1A1074696)

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Na-ju National Hospital. Depression [Internet]. Available: http://www.najumh.go.kr/html/content.do?depth=fc&menu_cd=02_02_00_01.
2. Organisation for Economic Co-operation and Development, A New Benchmark for Mental Health Systems: Tackling the Social and Economic Costs of Mental Ill-Health, Paris, France: OECD Publishing. 2021.
3. S.Y. Kim, J. E. Kim and H. W. Lee, “A Report of the Survey on the Current State of the Psychiatric Emergency Response in Seoul,” *The Mental Health*, vol. 11, no. 0, pp. 38-54, Dec. 2021. UCI: I410-ECN-0102-2022-500-001026410.
4. G. H. Song, M. K. Kim and A. S. Park, “Promising Services Based on AI for Mental Health,” *Electronics and Telecommunications Trends*, vol. 35, no. 6, pp. 12-23, Dec. 2020. DOI: 10.22648/ETRI.2020.J.350602.

5. M. Adam, M. Wessel and A. Benlian, "AI-based chatbots in customer service and their effects on user compliance," *Electron Markets*, vol. 31, no. 2, pp. 427-445, Mar. 2020. DOI: <https://doi.org/10.1007/s12525-020-00414-7>.
6. H. Rashkin, E. M. Smith, M. Li and Y. L. Boureau, "Towards Empathetic Open-domain Conversation Models: a New Benchmark and Dataset," in *Proceeding of the 57th Annual Meeting of the Association for Computational Linguistics*, Florence: Italy, pp. 5370-5381, 2019.
7. AI Hub. Wellness conversation dataset [Internet]. Available: <https://aihub.or.kr/opendata/kei-data/recognition-laguage/KETI-02-006>.
8. S. S. Kang, H. J. Won and M. H. Lee, "Analyzing Vocabulary Characteristics of Colloquial Style Corpus and Automatic Construction of Sentiment Lexicon," *Smart Media Journal*, vol. 9, no. 4, pp. 144-151, Dec. 2020. DOI: 10.30693/SMJ.2020.9.4.144.
9. K. R. Kim, J. H. Moon and U. R. Oh, "Analysis and Recognition of Depressive Emotion through NLP and Machine Learning," *The Journal of the Convergence on Culture Technology*, vol. 6, no. 2, pp. 449-454, May. 2020. DOI: 10.17703/JCCT.2020.6.2.449.
10. H. R. Seo and M. Song, "An Analysis of the Discourse Topics of Users who Exhibit Symptoms of Depression on Social Media," *Journal of the Korean Society for Information Management*, vol. 36, no. 4, pp. 207-226, Dec. 2019. DOI:10.3743/KOSIM.2019.36.4.207.
11. H. J. Chin, G. H. Baek, C. Y. Cha, J. H. Choi, H. S. Im and M. Y. Cha, "A study on the categories and characteristics of depressive moods in chatbot data", in *Proceeding of the Korea Information Processing Society Conference*, Yeosu: Korea, pp. 993-996, 2021.
12. Kind Korean Morpheme Analyzer [Internet]. Available: <http://kkma.snu.ac.kr/>.
13. T. Mikolov, K. Chen, G. Corrado and J. Dean, "Efficient Estimation of Word Representations in Vector Space," in *Proceeding of Workshop at International Conference on Learning Representations*, 2013.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.