A Deep Learning-based Depression Trend Analysis of Korean on Social Media*

딥러닝 기반 소셜미디어 한글 텍스트 우울 경향 분석

Seojeong Park (박서정)**, Soobin Lee (이수빈)***
Woo Jung Kim (김우정)****, Min Song (송민)*****

ABSTRACT

The number of depressed patients in Korea and around the world is rapidly increasing every year. However, most of the mentally ill patients are not aware that they are suffering from the disease, so adequate treatment is not being performed. If depressive symptoms are neglected, it can lead to suicide, anxiety, and other psychological problems. Therefore, early detection and treatment of depression are very important in improving mental health. To improve this problem, this study presented a deep learning-based depression tendency model using Korean social media text. After collecting data from Naver KonwledgeiN, Naver Blog, Hidoc, and Twitter, DSM-5 major depressive disorder diagnosis criteria were used to classify and annotate classes according to the number of depressive symptoms. Afterwards, TF-IDF analysis and simultaneous word analysis were performed to examine the characteristics of each class of the corpus constructed. In addition, word embedding, dictionary-based sentiment analysis, and LDA topic modeling were performed to generate a depression tendency classification model using various text features. Through this, the embedded text, sentiment score, and topic number for each document were calculated and used as text features. As a result, it was confirmed that the highest accuracy rate of 83.28% was achieved when the depression tendency was classified based on the KorBERT algorithm by combining both the emotional score and the topic of the document with the embedded text. This study establishes a classification model for Korean depression trends with improved performance using various text features, and detects potential depressive patients early among Korean online community users, enabling rapid treatment and prevention, thereby enabling the mental health of Korean society. It is significant in that it can help in promotion.

초 록

국내를 비롯하여 전 세계적으로 우울증 환자 수가 매년 증가하는 추세이다. 그러나 대다수의 정신질환 환자들은 자신이 질병을 앓고 있다는 사실을 인식하지 못해서 적절한 치료가 이루어지지 않고 있다. 우울 증상이 방치되면 자살과 불안, 기타 심리적인 문제로 발전될 수 있기에 우울증의 조기 발견과 치료는 정신건강 증진에 있어 매우 중요하다. 이러한 문제점을 개선하기 위해 본 연구에서는 한국어 소셜 미디어 텍스트를 활용한 딥러닝 기반의 우울 경향 모델을 제시하였다. 네이버 지식인, 네이버 블로그, 하이닥, 트위터에서 데이터 수집을 한 뒤 DSM-5 주요 우울 장애 진단 기준을 활용하여 우울 증상 개수에 따라 클래스를 구분하여 주석을 달았다. 이후 구축한 말뭉치의 클래스 별 특성을 살펴보고자 TF-IDF 분석과 동시 출현 단어 분석을 실시하였다. 또한, 다양한 텍스트 특징을 활용하여 우울 경향 분류 모델을 생성하기 위해 단어 임베딩과 사전 기반 감성 분석, LDA 토픽 모델링을 수행하였다. 이를 통해 문헌 별로 임베딩된 텍스트와 감성 점수, 토픽 번호를 산출하여 텍스트 특징으로 사용하였다. 그 결과 임베딩된 텍스트에 문서의 감성 점수와 토픽을 모두 결합하여 KorBERT 알고리즘을 기반으로 우울 경향을 분류하였을 때 가장 높은 정확률인 83.28%를 달성하는 것을 확인하였다. 본 연구는 다양한 텍스트 특징을 활용하여 보다 성능이 개선된 한국어 우울 경향 분류 모델을 구축함에 따라, 한국 온라인 커뮤니티 이용자 중 잠재적인 우울증 환자를 조기에 발견해 빠른 치료 및 예방이 가능하도록 하여 한국 사회의 정신건강 증진에 도움을 줄 수 있는 기반을 마련했다는 점에서 의의를 지닌다.

Keywords: topic modeling, deep learning, sentiment analysis, social media
토픽 모델링, 딥러닝, 감성분석, 소셜미디어

---

* This paper is based on a Master's Thesis of the Dept. of Library & Information Science, Yonsei University. This work was supported by a National Research Foundation of Korea grant funded by the Korean government (NRF-2018S1A3A2075114).
** Department of Library and Information Science, Yonsei University(pak9511@naver.com) (First Author)
*** Department of Library and Information Science, Yonsei University(bini122@yonsei.ac.kr) (Co Author)
**** Department of Psychiatry, Yongin Severance Hospital, Yonsei University College of Medicine, Yongin, Gyeonggi, Korea(woojungkim@yuhs.ac) (Co Author)
***** Department of Library and Information Science, Yonsei University(min.song@yonsei.ac.kr) (Corresponding Author)

■ 논문접수일자 : 2022년 2월 14일 ■ 최초심사일자 : 2022년 2월 28일 ■ 게재확정일자 : 2022년 3월 4일
■ 정보관리학회지, 39(1), 91-117, 2022. http://dx.doi.org/10.3743/KOSIM.2022.39.1.091

1. Introduction

Depression is a globally prevalent mental disorder, and according to the World Health Organization (WHO), 264 million people suffer from depression (WHO, 2020). However, in many countries, depression is not properly diagnosed and adequate treatment is not provided, leading to serious situations such as suicide (Friedrich, 2017; Tadesse, 2019). Accordingly, people use social media that allows them to talk about their own illness relatively freely even if they do not enter the hospital without time or space restrictions on their mental illness or depression symptoms. If people are diagnosed without the burden of time and money for their depressed state, and this makes it possible to detect depression early and lead to treatment, it will have a positive effect on people's mental health.

There are many attempts to perform Automatic Depression Detection (ADD) using social media data such as Twitter and Facebook, where people express and share their emotions in real-time (Guntuku et al., 2017). However, the previous studies based on social media mainly analyzed data in English. They are not compatible with the domestic environment because they have mainly conducted studies based on English data. This has the advantage of being able to focus more on the degree of depression tendency that exists in Korean society. Therefore, to tackle this critical limitation of the previous studies, in this study, data was collected on social media texts in Korean, and a corpus was constructed by annotating the degree of depression in each document. At this time, based on the guidelines verified by psychiatrists, the literature was annotated to establish a corpus with a reliable diagnosis level. This study proposes a deep learning-based depression tendency classification model to detect depression efficiently. In addition, to improve the performance of the depression tendency classification model, we utilized the embedded texts, emotional scores, and topics of the literature as various text features for elaborate classification of depression trends.

This study helped lay the foundation for mental health in Korean society by early detection of potential depressed patients among Korean online community users, enabling rapid treatment and prevention. The research questions set up in this study are as follows. First, we examined whether differences and characteristics are there between classes of Korean corpus. Second, we proposed the deep learning-based depression tendency classification model combining text features perform better than the existing text-based classification model. Third, we investigated whether text features improve the performance of the depression tendency classification model.

To this end, this study intends to use the Mental Illness Diagnosis Manual DSM-5 provided by the American Psychiatric Association from four social media data sources to annotate the degree of depression by class. Afterwards, to examine the characteristics of each class, TF-IDF analysis and simultaneous occurrence word analysis are performed, and word embedding, dictionary-based emotion analysis, and LDA topic modeling are performed to generate a depression tendency classification model using text characteristics.

When a depression tendency model using various text features, examined which text features have an impact on performance.

The rest of the paper is organized as follows. We will look at studies on the detection of depression in social media and previous studies on text feature extraction. In the methodology, the Korean corpus construction process will be introduced, and a depression tendency classification model using this corpus will be presented. In the conclusion, we will look at the contributions of the study.

2. Related work

2.1 Detection of depression in Social media

Automatic depression detection on social media detects and predicts people with depressive symptoms through observation of large-scale data (Guntuku et al., 2017). To this end, the collection of data related to depression in social media can be divided into four methods. The first is a method of collecting social media texts of users diagnosed with depression through self-assessment scales such as CES-D and BDI-II. Tsugawa et al.(2015) suggest users' depression was discriminated against using self-assessment scales, CES-D and BDI, and if the score between the two scales showed a low correlation, it was considered unreliable. As a result, a machine learning-based depression prediction model was created by collecting data such as the text of 209 tweets, the number of followers, and the number of tweets posted during the day. The second is to collect social media data that reveals whether you have diagnosed depression. Resnik et al.(2013) and Coppersmith et al.(2014) used tweets describing one's diagnosis of depression, such as "I was diagnosed with depression today" were extracted and used as a dataset. Based on this, an elaborate topic model was created to explore its potential effectiveness in automatic depression identification. Third, there is a way to collect data from online forums and related websites, which are spaces for discussions about depression. De Choudhury et al.(2016) used Reddit as one of the online forums. In this study, subreddit data related to mental disorders such as 'depression,' 'bipolar disorder,' 'eating disorder,' 'self-harm,' and 'post-traumatic stress disorder' were collected, and among users who uploaded posts to the subreddit, later 'Suicide Users who uploaded posts were also identified on the Watch' subreddit. Using this, we created a classification model that identifies users who have posted on mental health concerns which-change into discussions about suicide. Finally, based on social media data including keywords related to depression, researchers can collect data by annotating depression status according to guidelines. Mowery et al.(2017) collected tweets using keywords related to depression suggested by clinical psychologists, and three researchers annotated which symptoms of the DSM-5 major depressive disorder diagnosis criteria corresponded to the text. They created a machine learning-based depression detection model to determine the presence or absence of depressive symptoms using the annotated data, and additionally, data with depressive symptoms were three types of ‘depressed mood,’ ‘sleepless,’ and ‘fatigue or energy loss’ A model for classifying symptoms was created.

In this study, the burden of time and economic costs was reduced by collecting data from social media that can collect more data than a study using a survey. In addition, for the objective diagnosis of depression tendency according to medical guidelines, an annotation method was selected based on the DSM-5 diagnostic manual of the American Psychiatric Association. This is a complementary method to shed light on the living conditions associated with depression that are not captured by traditional depression diagnostic criteria (Guntuku et al., 2017).

Various machine learning and deep learning-based algorithm classifiers were used for depression detection and prediction. As a study using deep learning algorithms for depression detection and prediction, there is a study of Orabi et al.(2018). They performed depression detection by collecting Twitter texts from users with depression and users without depression. First, the text was embedded using the Word2Vec model, and depression detection was attempted by using deep learning algorithms such as Bidirectional LSTM with attention, a CNN, and RNN-based algorithm based on the embedded text. Compared with the performance of the SVM-based depression classification model, which is a machine learning algorithm, the performance of the classification model using the CNN algorithm is significantly improved, and it is found that the deep learning algorithm is suitable for depression detection. Therefore, in this study, depression detection was attempted using a deep learning algorithm that was able to understand the meaning inherent in the literature and solve the difficulty of extracting features for new literature based on a pre-learned model. Also, since many studies have attempted to detect depression using English data, studies on the detection of depression in social media using Korean data are rare. Therefore, in this study, we collect Korean social media data and propose a depression detection method suitable for Korean characteristics based on this data.

2.2 Text feature extraction

Text feature extraction is to extract information that can represent text from text, and through this, the performance of the learning algorithm can be improved and the learning time can be shortened (Liang et al., 2017). The purpose of text feature extraction is to represent a document as a multidimensional vector. For this, text feature extraction methods such as Document Frequency and TF-IDF(Term Frequency-Inverse Document Frequency) have been widely used (Cheng & Chen, 2019). However, traditional text feature extraction methods using the frequency of occurrence of words such as BOW (Bag of Words) have a limitation in that they cannot reflect the semantic relationship between words (Ruas et al., 2020). Therefore, the BERT model, which can capture the syntactic and semantic aspects of words, is used as an alternative to text extraction methods (Chronis & Erk, 2020).

In order to improve the performance, studies were conducted in which a classification model was trained by combining various text features. These include studies that improve performance by combining text features extracted through the word embedding method (Lilleberg, Zhu, & Zhang, 2015; Pasupa & Ayutthaya, 2019). In the study of Pasupa and Ayutthaya (2019), word embedding, part-of-speech tags, and sentiment vectors were used as text features to classify Thai sentiment. These text features were combined and used as input values for deep learning algorithms CNN, LSTM, and Bidirectional LSTM. In the results of the emotional classification experiment, the highest performance was achieved when CNN was used as the classification algorithm and all three text features were used for classification. Therefore, this study aims to improve the performance of the deep learning-based depression tendency classification model and shorten the training time by combining various text features such as emotion scores and topics to the embedded text.

3. Method

3.1 Developing annotated corpus

In this study, social media texts in Korean were collected to build a Korean corpus for depression tendency classification. The collected data were manually annotated by the researchers to classify depression tendency classes. To classify depressive tendencies, we need a labeled class, but there are no previously published indicators or corpus, so we created it ourselves. Unlike previous studies, in this study, it is possible to present the diversity of the depression tendency corpus by constructing the Korean corpus, and it is possible to measure the local depression tendency index concentrated in Korean society, thus connecting the social atmosphere and the corpus. You can build a foundation for thinking. The production of the corpus consisted of two steps. First, data was collected from Naver KnowldegeiN, Hidoc, Naver Blog, and Twitter, and then the researchers directly annotated the data.

3.1.1 Data collection

In this study, four social media data sources were used to collect data for annotation. Data was collected from Naver KnowledgeiN, Hidoc, Naver Blog, and Twitter, which have the aforementioned social media characteristics. Naver was founded in 1999 as the most popular search engine in Korea. Therefore, since 2002, Naver added the utility of search by creating a community that answers questions and answers like Naver KnowledgeiN (Nam, Ackerman, & Adamic, 2009). Naver KnowledgeiN often ask realistic questions, and the answers are reliable in terms of the fact that when a user asks a question, an expert answers it. Hidoc is a site that has partnered with Naver KnowledgeiN and provides counseling services with 2,500 specialists participating. Hidoc differs from Naver KnowledgeiN in that it is a community that provides health-specific questions and answers that can check health information through columns or articles and exchange health-related questions. Data collection through the Hidoc site, where communication focused on health, is easy to determine the degree of depression tendency, so a Hidoc data source was selected. Naver Blog is the largest blogging service in Korea, and it allows bloggers to communicate their opinions and experiences to others on a topic (Moon & Han, 2011). Therefore, users can share informational articles about emotions or depression symptoms they feel in their daily lives through blogs. Twitter is a data source mainly used in studies on depression in social media, and is widely used in that it posts content on user behavior, activities, thoughts, and emotions that can indicate emotional states in public spaces. (Conway & O'Connor, 2016). For this reason, data from January 2017 to December 2019 were collected with the search term 'depression(우울증)' from four selected data sources.

As for the collection results by data source, a total of 442,859 data were collected as shown in <Table 1>. In the case of Naver KnowledgeiN and Hidoc, only user questions were collected.

3.1.2 Data annotation

Among the collected 442,859 data, 4,604 data were randomly extracted from each of the four data sources using a random sampling method and the depression tendency annotation was performed. As a criterion for diagnosing depression tendency, the criteria for diagnosis of depression shown in the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) were used. The criteria for diagnosis of depressive disorder indicated in DSM-5 include at least five symptoms: depressed mood, decreased interest, weight change, sleep change, psychomotor agitation or delay, energy loss, sense of worthlessness or guilt, decreased attention, suicidal thoughts and attempts. Among the five or more diagnostic criteria, one or more symptoms of depression and decreased interest should be present.

However, unlike in clinical settings where a doctor examines all symptoms through an interview and diagnoses depression, in this study, three classes were classified in that the degree of depression tendency was judged based on symptoms written in one document. In order to meet the DSM-5 Depressive Disorder Diagnosis Criteria, five or more symptoms must be met. However, on social media, the author who writes does not always describe all symptoms and circumstances of himself or others. As there is a point that the diagnosis of depressive disorder may not be professional or rigorous, the slightly relaxed class criteria were applied as shown in <Table 2>. When applying the class classification criteria, it was verified by a psychiatrist.

〈Table 1〉 Number of data collected from 4 data sources

Data Source	Number of Data	Number of Annotation Data
Naver KnowledgeiN	224,172	1,461
Naver Blog	158,739	1,488
Twitter	58,873	1,329
Hidoc	1,075	326
Total	442,859	4,604

〈Table 2〉 Class separation criteria

Class number	Class name	Class criteria
0	Non-depressed literature	Hospital PR articles, Advertising articles, News articles, reviews etc.
1	Latent-depressed literature	Be relevant depressed mood or lack of interest. Not more than one other symptom of depression.
2	Depressed tendency literature	Be relevant depressed mood or lack of interest. More than two other symptoms of depression.

A total of 8 researchers annotated 4,604 data records collected according to the validated guidelines. These eight researchers consisted of six master's students and two doctoral students majoring in informatics and were familiar with the DSM-5 diagnostic criteria for depression. The annotation was carried out twice, and the first annotated data was cross-validated twice. Documents that were not consensus in cross-validation were discussed and concluded. In addition, some data among the annotation results were verified by a psychiatrist to perform accurate annotation. Out of 4,604, the 50 data records were the most difficult to agree on during the annotation process. Of the 50 data records, the annotation results of 37 data records were consistent (74%). The data records with disagreement were further discussed by all participating researchers and drew the final consensus.

As a result of data annotation, the number of non-depressed documents with no depression tendency at all was 1,908, and 1,642 documents with latent depression, a class with a potential depression tendency. The number of documents with a tendency toward depression was 1,054.

When explaining the data annotation process through the examples of documents for each class in <Table 4>, the first document was classified as Class 1, which is a non-depressing document because it was the contents of reading a book and writing a remark. The second document indicated that there was depression, and marked it with one annotation, and because there is an expression "the head was not recognized", the annotation was marked again. This can be classified into class 2, which is a latent depressive document, because there is one depressive mood symptom and one other depressive symptom. The third document indicated that he was taking drugs for depression as one annotation, annotated anxiety, and annotated self-harm and lack of sleep. This can

〈Table 3〉 Data annotation result

Class number	Class name	Number of data
0	Non-depressed literature	1,908
1	Latent-depressed literature	1,642
2	Depressed tendency literature	1,054

〈Table 4〉 Data annotation example

Literature example	Class number
A phrase from Freud's "Morning and Depression" ("Sad and Depression" in Volume 1 of the Complete Book of Open Books); Why is there no sign of creating its own economic conditions for the moment of victory even after the grief is over? After reading these verses, I feel depressed as well. A mourning without weakness... I remember the families of the Sewol ferry. 프로이트의 '애도와 우울'(열린 책들 전집판1 1권 중 '슬픔과 우울증')의 한 구절: 어째서 슬픔이 다 끝난 뒤에도 승리의 순간을 위해 그 나름의 경제적 조건을 조성하는 조짐이 없는 것일까? 이런 구절을 읽고 나면 나도 마냥 우울해진다. 기약 없는 애도라니... 세월호 가족들이 떠오른다.	0
When I was seriously **depressed...** I went to see an art exhibition while feeling emotional. I couldn't read the text there. It was written in large letters, but no matter how much I looked at it, **I couldn't recognize it in my head..** Among the symptoms of depression I knew there was a decrease in brain function, but it was a very, very shocking experience. 우울증 심각했을 때.. 젤 충격이었던 경험은 기분 환기도 할 겸 미술전시 보러 갔는데 내가 거기서 글을 못 읽는 거였다.. 큰 글씨로 작품설명 써져 있는데 아무리 쳐다봐도 그게 **머리에 인식이 안됐다..** 우울증 증상 중 뇌기능 저하 있는 건 알았지만, 직접 겪으니 너무너무 충격이었음.	1
I'm taking medicine **for depression...** I'm a little scared of people, I'm afraid of eye contact, I'm afraid that someone will swear me from behind and ignore me because **I'm afraid, I'm anxious and afraid.** Even if I'm fine, if I meet people and come home, I feel upset and very bad. **Anxious, anxious** and crazy fear because of it. When I get too depressed, sad, lethargic, and even worse (if the other person seems to be ignoring me at all), **I even hurt myself.** With the knife... it's painful. I feel like I'm getting a little better if I feel like I'm going to get back to myself by abusing myself. If the other person seems to be ignoring me a little, it's hard to bear it. It's hard to live because I care too much about being a person. It's hard to walk outside because I hate people and I'm scared. I don't go out if it's an important task. I don't know why I am doing this. It is this if you only meet people. People who pass by on the street don't know, so it's not too bad. When I meet people I see or know, I feel so anxious that I'm going crazy, and I **can't sleep** because I care about it. […] 우울증으로 약 복용중인데요... 제가 사람이 좀 무서워요 눈 마주치기도 무섭고 또 누가나를 뒤에서 욕할까봐 날 안 좋게 볼까봐 무시할까봐 **걱정되고 불안하고 무섭고** 이런 증세가 있거든요. 잘 있다가도 사람을 만나고 오면 집에 와서도 기분이 찝찝하고 너무 안 좋아요. 그것 때문에 **불안 초조하고** 미칠 것 같은 두려움. 너무 우울 슬퍼지고 무기력해지고 심지어 더 심할 때는 (상대방이 나를 조금이라도 무시하는거 같아 보이면) **자해까지합니다.** 칼로.. 괴로워서요. 나자신을 학대해서 내 스스로 정신에 나사 돌아갈 것 같아져야 조금이라도 나아지는거 같아서요. 상대방이 조금만 절 무시하는 거 같아 보이면 그만큼 견디기 힘듭니다. 사람이란게 너무 신경쓰여서 살기 힘들어요. 사람싫고 무서워서 밖에 돌아다니기도 힘들어요. 중요한 업무 아니면 안 나갑니다. 대체 제가왜 이러는지 모르겠어요. 사람만 만나고 오면 이럽니다. 길거리 지나가는 사람들은 모르는 사람이니 많이 심하지 않는데요. 자주보는 사람이나 아는 사람을 만나면 너무 불안해서 미칠거 같고, 신경 쓰는 나머지 **잠도 잘 안 옵니다.** […]	2

be classified into class 3, which is a depressive tendency document, because it is included in two or more other depressive symptoms. We decided to publish the entire data tagged in this way to the address we wrote in the appendix.

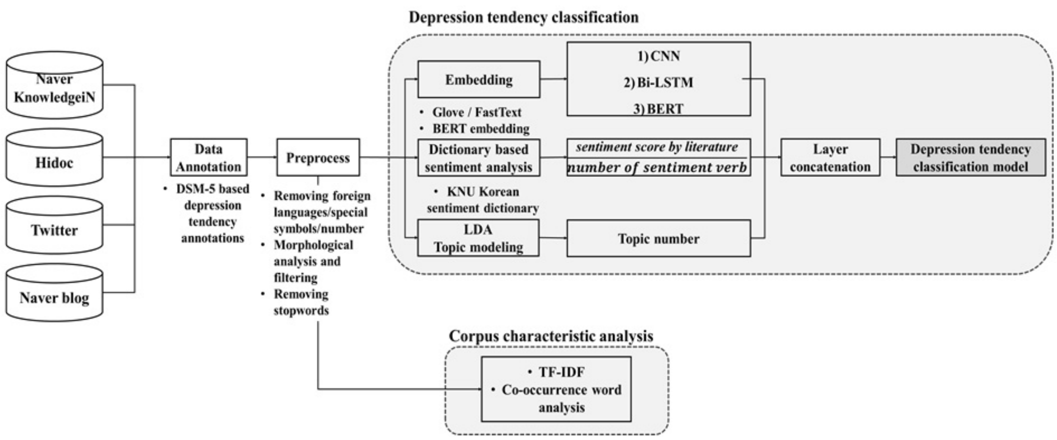
3.2 Research process

After 4,604 annotating data from four Korean social media data sources, pre-processing such as morpheme analysis and stopword removal was performed on the corpus constructed based on the DSM-5 standard. Subsequently, in this study, TF-IDF analysis and word occurrence analysis were performed for each corpus class to understand the differences between classes and the characteristics of each class.

To build a depression tendency classification model, pre-processed data was used to perform pre-based emotional analysis and LDA topic modeling, and this was used as a text feature. Emotional analysis was performed under the assumption that texts containing depression signs may contain more negative emotions than texts that do not (Trotzek, Koitka, & Friedrich, 2018), and emotional scores for each document were used as text features. In addition, the topic model can effectively classify topics by summarizing a large amount of literature and predicting features in the literature, and topics calculated through LDA topic modeling have also been used as a feature of the classification model. Accordingly, the topic was used as a text feature by referring to studies with improved performance (Zhang et al., 2014; Resnik, Garron, & Resnik, 2013) by adding a topic as a feature. In this study, topic numbers for each document, calculated using LDA topic modeling, were used as text features. After that, a depression tendency classification model was created by combining the previously calculated text features and another text feature, the embedded text. The overall research method and process is shown in Figure 1.

3.2.1 Preprocess

In this study, text pre-processing was performed before corpus characteristic analysis and depression tendency classification model were generated. Foreign languages, special symbols, and numbers were removed because the study was conducted in Korean. Afterward, morphological analysis was performed using Mecab. At this time, the user dictionary was used to solve the case where various nouns such as drug names and new words cannot be properly analyzed due to the finiteness of words contained in the corpus of the morpheme analyzer. In the user dictionary,



〈Figure 1〉 Research flow

drug names such as 'Belproic acid,' 'Lamotrigine,' and 'Benzodiazepine' were included, and health foods such as 'Moringa' and 'Yellow Mushroom' were also included. In addition, disease names and symptoms such as 'schizophrenia' and 'abdominal distension' were included, and new words such as 'chwijun' were also included.

Afterward, the tokenized words were removed with a morpheme analyzer. Stop words include words that are collected regardless of the content of the text when collecting data, such as 'neighbor,' 'font,' and 'copy,' and meaningless connection words such as 'and' and 'such'. After that, only words corresponding to nouns were left, and TF-IDF analysis and simultaneous word analysis were performed, and all parts of speech except stop words were used in the generation of the depression tendency classification model.

3.2.2 TF-IDF and Word co-occurrence analysis

TF-IDF (Term Frequency-Inverse Document Frequency) means the product of TF and IDF, and TF is a measure of the frequency of a term in a document or a set of documents (Aizawa, 2003). IDF is a measure of the frequency of inverse words in the entire literature (Yun-tao, Ling, & Yong-cheng, 2005). TF-IDF is a numerical statistic that can represent keywords in specific literature, and these keywords can be categorized or identified (Qaiser & Ali, 2018). In this study, the TF-IDF values of the top 5 words were calculated for each document of the constructed corpus. After that, each class was sorted in descending order based on the TF-IDF value, and the top 30 words were compared. Through this, we tried to analyze the characteristics and differences of each class.

Simultaneous word analysis is a content analysis technique that is useful for mapping information similarity between items in text (Callon et al., 1983; Wang et al., 2012). After analyzing the words that appear simultaneously in one document, the concurrent appearance network was visualized using Gephi, a data visualization tool. In a co-occurrence network, a node represents a word and an edge, a line connecting nodes, represents the frequency of simultaneous appearance. A group of words with frequent simultaneous occurrences was analyzed using a community detection algorithm (Blondel et al., 2008) that divides frequently occurring words into a group. In addition, to determine which node is the most important in the entire network, the degree centrality was measured and the frequency of simultaneous occurrence was used as the size of the node. In this study, we tried to analyze the characteristics and differences of each class by creating a simultaneous appearance network for each class in the same way as above.

3.2.3 Feature extraction

In this study, a depression tendency classification model is created by using embedded text, emotion scores, and topics as text features. Glove and FastText word embedding models in addition to BERT pretrained model were used for document embedding, and dictionary-based sentiment analysis and LDA topic modeling were used to calculate the emotion score and topic information of the document.

Word embedding has utility that plays a key role in fundamental problems in many NLP (Natural Language Processing) fields (Turney & Pantel, 2010; Yin & Shen, 2018) and is a vector space for representing the meaning of words (Zhao et al., 2018). In this study, Glove and FastText word embedding models were used to use CNN and Bidirectional LSTM deep learning algorithms. Glove is a global log-bilinear regression model used to express unsupervised learning words, and shows the superior performance when performing word analysis and word similarity analysis entity name recognition compared to other models (Pennington, Socher, & Manning, 2014). FastText has the advantage that each word is a sum of vectors and each vector represents an n-gram. This approach has the advantage of a learning process that emphasizes the common roots of words (Athiwaratkun, Wilson, & Anandkumar, 2018). In addition, when a pre-trained model-based word embedding was performed, a pre-trained BERT (Bidirectional Encoder Representations from Transformers) word embedding model was used. BERT is a model trained to output embedding vectors of a specific word according to a given context (Devlin et al., 2018). In this study, the document was embedded using the above word embedding models, which were fed as the input for the deep learning algorithm.

Dictionary-based sentiment analysis is a technique in which a researcher builds a sentiment dictionary composed of polarities that indicate the degree of positive and negative words, and uses it to perform sentiment analysis. In this study, emotion analysis was conducted using the KNU Korean Emotion Dictionary (KNU Korean Emotion Dictionary, 2018). The KNU Korean Sentiment Dictionary is a general-purpose sensibility dictionary that can be used in various fields, and includes the sensibility degree of vocabulary in various forms. In the dictionary, a unigram consisting of a single word, a bigram that is a combination of two consecutive words, a phrase, a sentence pattern, an emoticon, etc. are calculated. The level of sensitivity of each vocabulary is divided into 'very negative (-2),' 'negative (-1),' 'neutral (0),' 'positive (1),' and 'very positive (2)'.

In this study, using the KNU Korean Sentiment Dictionary, the values of the level of emotional vocabulary appearing in one document as shown in Equation (1) are summed, and the average value divided by the number of emotional vocabulary appearing in one document is used as the emotional score for each document (Lalithamani, Thati, & Adhikesavan, 2014). This is to average and use it as a text feature because each document has a different length.

$$Sentiment\ Score = \frac{\sum_{i=1}^{n} Polarity\ Socre(T_i)}{n} \tag{1}$$

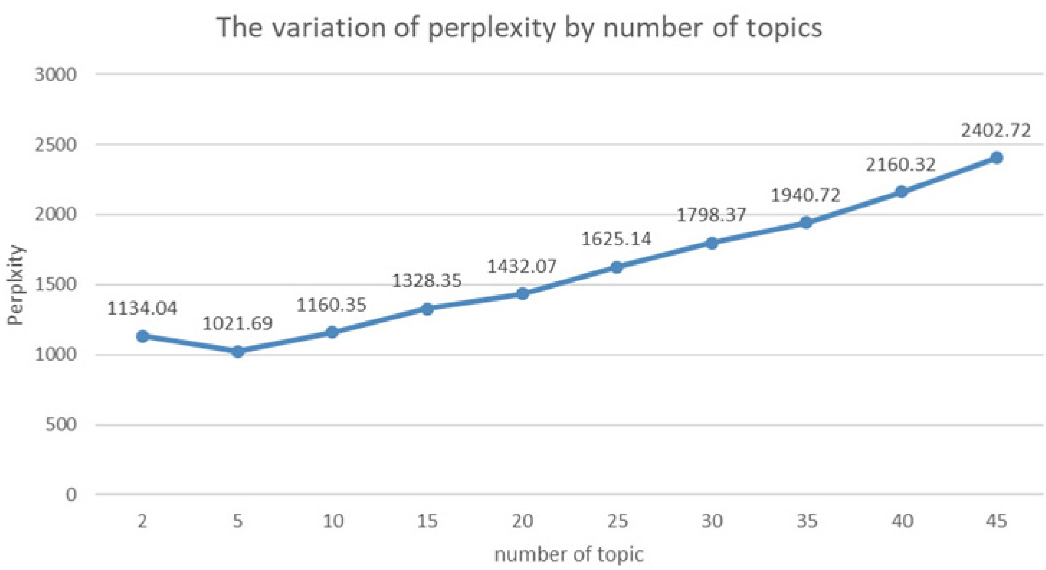
In Equation (1), n denotes the number of emotional words appearing in the document, and denotes the i-th emotional vocabulary in the document. The polarity score means the positive and negative degree values of the emotional vocabulary, and the emotional score means the emotional score of the literature. The emotional score calculated through Equation (1) is expressed as one emotional score per document, which is used as an input value of the depression tendency classification model.

In this study, LDA (Latent Dirichlet Allocation) topic modeling (Blei, Ng, & Jordan, 2003) was performed to understand what topics each document is discussing. LDA assigns the subject of literature and generates the distribution of the subject of words in a literature group (Patterson et al., 2010). For the constructed corpus, the number of topics starts from 2 and increases by 5 units such as 5, 10, and 15, and the number of topics with the lowest perplexity is selected by experimenting up to 45. The topic of was calculated. The change in the degree of confusion according to the number of topics is shown in Figure 2, and since the degree of confusion is the lowest when the number of topics is 5, the topics for each document of the previously constructed corpus were calculated based on five topics.

3.2.4 Deep Learning-based text classification using additional features

In this study, various text features were extracted through the method described above, and emotional scores and topics were combined to classify depression trends. To this end, text features are learned using CNNs (Convolution Neural Networks) or Bi-LSTM (Bi directional Long Short Term Memory) algorithms. The CNNs algorithm was originally developed for computer image processing, but as it was known to be effective for text classification, one of the natural language processing tasks, a sentence classification method using word embedding and CNNs was proposed, which showed excellent performance (Kim, 2014).

However, studies that classified text using CNNs have a limitation in that they cannot provide performance when dealing with long sentences that have a



<Figure 2> The variation of perplexity by number of topics

complex semantic relationship with the whole text. RNNs (Recurrent Neural Networks) were developed to compensate for these limitations, but there are drawbacks in that gradient loss or explosion problems occur and long-term dependence problems (Liu & Guo, 2019). To overcome this, a bidirectional LSTM combining bidirectional RNN and LSTM was proposed. Bi-LSTM is an additional development of LSTM, which allows access to content by combining forward and reverse hidden layers (Liu & Guo, 2019). Therefore, in this study, the depression tendency classification was performed using the CNN deep learning algorithm, which is known to be effective for text classification, and the Bidirectional LSTM deep learning algorithm, which supplemented the limitations of CNN and RNN.

In addition, in this study, BERT was selected as a deep learning algorithm to classify depression trends. After BERT pre-trained to learn effective embeddings with a large number of documents, the trained model achieved good performance in various natural language processing tasks such as question-and-answer through fine-tuning. BERT is significant in that it showed good performance for various natural language processing tasks simply by fine-tuning by stacking one more layer while maintaining the structure of the pretrained model. For processing Korean, you can use the multilingual BERT, which was first released by Google. This is a multilingual model that has been pre-trained on the Wikipedia corpus of 104 languages, and is a universal model that can cope with multiple languages. Multilingual BERT showed good results in the natural language processing task, but did not achieve good performance for languages other than English (Martin et al., 2019). In order to overcome these limitations, the Electronics and Telecommunications Research Institute (ETRI) released the Korean deep learning language model KorBERT. Unlike English, which is an inflectional language whose ending changes depending on its application, Korean is an abbreviated word that forms a single word by combining content words (nouns/verbs) and functional words (investigations/endings). Tokenization is required. KorBERT collected Korean encyclopedic texts and Korean newspaper articles for about 15 years, and learned 4.7 billion morphemes for 23 GB of text. As a result of conducting natural language processing tasks such as classifying document subjects using the model, it showed better performance than Google's multilingual BERT (Lim, Kim, & Kim, 2020). Therefore, in this study, two BERT models were used to classify depression trends and compare the results.

4. Results

In the research results, we present the results of the analysis of TF-IDF and simultaneous occurrence words used to investigate the characteristics of the constructed corpus. Also, using the text extraction method and deep learning algorithm described above, we intend to implement a depression tendency classification model utilizing various text features. Through the experiment of the depression tendency classification model, we would like to see which text features are used to improve performance.

4.1 TF-IDF analysis by corpus class

In this study, TF-IDF analysis was performed to examine the characteristics of each class of corpus training the depression tendency classification model. The subjects of analysis were document groups by class, and the documents were divided by class, and the TF-IDF value of each document was calculated, and the top 30 words of the TF-IDF value were sorted. The results of the top 30 TF-IDF values in non-depressed literature are shown in the table below.

Words such as ‘saffron,’ ‘moringa,’ ‘ulgeum,’ ‘omega,’ ‘cheonma,’ and ‘vitamin’ in the non-depressive literature were related to health foods known to relieve depressive symptoms. This is interpreted as the fact that the non-depressing literature mainly consists of informational posts and advertisements, so informational posts and advertisements related to health foods ranked high in TF-IDF. In addition, words such as ‘hives,’ ‘dizziness,’ ‘pain,’ ‘thyroid,’ and ‘menopausal’ appeared in the literature mainly when describing side effects information of depression-related drugs or information related to causes and symptoms of depression. Words such as ‘the attending physician,’ ‘pilates,’ ‘coaching,’ and ‘hypnosis’ appeared in informational posts describing ways to alleviate and treat depressive symptoms, or in posts asking how to relieve depressive symptoms.

Words such as ‘weather,’ ‘caffeine,’ and ‘home’ in the latent depression literature class are described as factors that influence depression in the literature. ‘Depression,’ ‘degeneration,’ ‘motivation,’ ‘memory,’ ‘meaning,’ ‘conversation,’ ‘powerlessness,’ and ‘illustration’ were words that appeared in the literature to

〈Table 5〉 Top 30 words of TF-IDF value in Non-depressed literature

Rank	Term	TF-IDF value	Rank	Term	TF-IDF value
1	Dobermann(도베르만)	0.990986	16	Turmeric(울금)	0.88209
2	Saffron(사프란)	0.929592	17	Omega(오메가)	0.880387
3	Lack(부족)	0.925721	18	Signal(신호)	0.875182
4	Insurance(보험)	0.916557	19	Rash(두드러기)	0.874703
5	Danger(위험)	0.915136	20	Gastrodia(천마)	0.870077
6	Friend(친구)	0.912185	21	Menopause(갱년기)	0.867416
7	Thyroid(갑상선)	0.901443	22	Firefly(개똥벌레)	0.865343
8	Ring(반지)	0.900151	23	Wife(아내)	0.864944
9	Spouse(배우자)	0.900005	24	Dizziness(어지럼증)	0.864759
10	Teacher(강사)	0.899358	25	Sherlock(셜록)	0.862009
11	Lamp(조명)	0.894052	26	Hug(허그)	0.860424
12	Moringa(모링가)	0.893711	27	Vitamin(비타민)	0.86028
13	Pain(통증)	0.893616	28	Coaching(코칭)	0.859832
14	family doctor(주치의)	0.893591	29	Hypnosis(최면)	0.858622
15	Pilates(필라테스)	0.890362	30	Euthanasia(안락사)	0.856829

〈Table 6〉 Top 30 words of TF-IDF value in Latent depressed literature

Rank	Term	TF-IDF value	Rank	Term	TF-IDF value
1	Weather(날씨)	0.986905	16	Aunt(숙모)	0.883589
2	f**k(시*)	0.986905	17	Family(가정)	0.876959
3	Caffeine(카페인)	0.986188	18	Control(조절)	0.863449
4	Bicycle(자전거)	0.984557	19	adult disease(성인병)	0.861702
5	Depression(우울)	0.984423	20	Limit(극한)	0.859645
6	Fool(바보)	0.981947	21	Atrophy(퇴화)	0.858177
7	Person(인간)	0.968367	22	Mother(엄마)	0.856514
8	Will(의욕)	0.960831	23	emergency room(응급실)	0.854978
9	Reason(이유)	0.958255	24	Cure(치유)	0.853784
10	Healing(힐링)	0.945637	25	Treatment(치료법)	0.853474
11	Memory(기억력)	0.944669	26	Surprise(놀람)	0.850778
12	Mind(정신)	0.921547	27	Lethargy(무기력)	0.846443
13	Meaning(의미)	0.906088	28	Smoking(흡연)	0.841284
14	Comfort(위로)	0.886808	29	Episode(삽화)	0.832873
15	Conversation(회화)	0.885459	30	Exercise(운동)	0.832641

〈Table 7〉 Top 30 words of TF-IDF value in Depressed tendency literature

Rank	Term	TF-IDF value	Rank	Term	TF-IDF value
1	Menopause(갱년기)	0.872203	16	Father(아버지)	0.672883
2	Division(분열)	0.870364	17	Pain(통증)	0.670561
3	self-injury(자해)	0.856087	18	Bipolar(양극)	0.664753
4	Officer(장교)	0.849427	19	Husband(남편)	0.663818
5	God(하나님)	0.805461	20	Chronicity(만사)	0.661205
6	Menses(생리)	0.798976	21	a struggle with disease(투병)	0.657363
7	Cousin(사촌)	0.787402	22	Wife(아내)	0.65169
8	Mother(어머니)	0.763979	23	School(학교)	0.651282
9	Drowsiness(졸음)	0.760216	24	law school(로스쿨)	0.649366
10	Mom(엄마)	0.738679	25	fake illness(꾀병)	0.647866
11	Try(시도)	0.694452	26	Gene(유전자)	0.647346
12	Alzheimer(치매)	0.685206	27	Syndrome(증후군)	0.644192
13	Bronchitis(기관지염)	0.683215	28	Schizophrenia(조현병)	0.637982
14	Mood(기분)	0.682196	29	Crying(울음)	0.637394
15	Weight(몸무게)	0.678415	30	Burnout(번아웃)	0.634165

describe the symptoms of depression in oneself or others. Words such as 'healing,' 'comfort,' 'cure,' and 'exercise' are described in the context of trying to overcome the symptoms of depression.

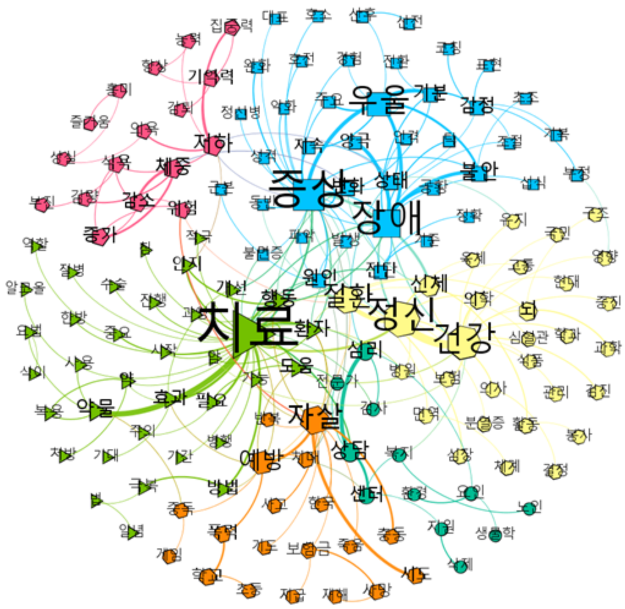
'Division' in the depressive tendency literature class is a word related to schizophrenia, and 'bipolar' is a word referring to bipolar affective disorder. In other words, the literature on the tendency of depression was composed of words related to mental disorders such as 'division,' 'bipolar,' and 'schizophrenia’. Words related to depressive symptoms such as ‘self-harm,’ ‘drowsy,’ and ‘burnout’ also ranked above the TF-IDF value.

In summary, what each class has in common is that words related to depression (Schwartz et al., 2014) that occur during a specific period, such as ‘menopausal’ and ‘menstruation,’ were located at the top of TF-IDF. On the other hand, each class showed a difference in words related to depressive symptoms relief and treatment, such as ‘the attending physician’ and ‘the treatment method’. In the non-depressive literature class, words related to depressive symptom relief and treatment accounted for one third of the top 30 words. In the Latent Depression Literature Class, 5 of the top 30 words were related to depression symptoms relief and treatment. However, the word corresponding to this did not exist at the top in the depressive tendency literature. Words such as ‘cousin,’ ‘mother,’ ‘mom,’ and ‘father’ were also placed at the top.

4.2 Co-occurrence word analysis by corpus class

In this study, co-occurrence word analysis was conducted to examine the overall trend of discussions related to depression in which words in each class of corpus. The analysis target is the top 500 word pairs obtained based on the number of simultaneous appearances in each class data set of the corpus. Based on this, the simultaneous appearance network was visualized. At this time, after filtering with the Giant component, we tried to detect the community by using the community algorithm method, Modularity. Based on the results of modularity calculation, words that frequently appear simultaneously in the network are grouped into a group and displayed in the same color and shape. For effective visualization, communities with less than 3.0% modularity were excluded from visualization. We used DSM-5 to analyze and visualize co-occurrence word analysis by dividing them into three classes, annotated and classified, such as ‘non-depressive literature,’ ‘latent depressive literature,’ and ‘high depression tendency literature’.

The co-occurrence word network of the non-depressive literature class includes 155 nodes and 232 edges, and consists of a total of 6 communities. The result of visualizing the simultaneous appearance network of the non-depressed literature class is shown in <Figure 3>, and the words belonging to each community are shown in <Table 8>. As a result, the word with the greatest connection centrality in the entire network was ‘treatment,’ followed by words such as ‘symptoms,’ ‘health,’ ‘disability,’ and ‘spirit’. In addition, ‘prevention,’ ‘counseling,’ and ‘drug’ are words that are located above the center of connection. From this, it can be seen that keywords related to the treatment of depression are important in the network of simultaneous appearance of non-depressive literature. The word network for simultaneous occurrence of non-depressive literature includes community 0 including words such as ‘treatment,’ ‘taking,’ and ‘overcoming,’ and community 4 including ‘doctor’ and ‘examination,’ and ‘counseling’ and ‘psychology’. From the included community 5, it can be seen that contents related to overcoming depression and treatment methods are dominated. In addition, there was a community of



<Figure 3> Co-word network in Non depression related texts

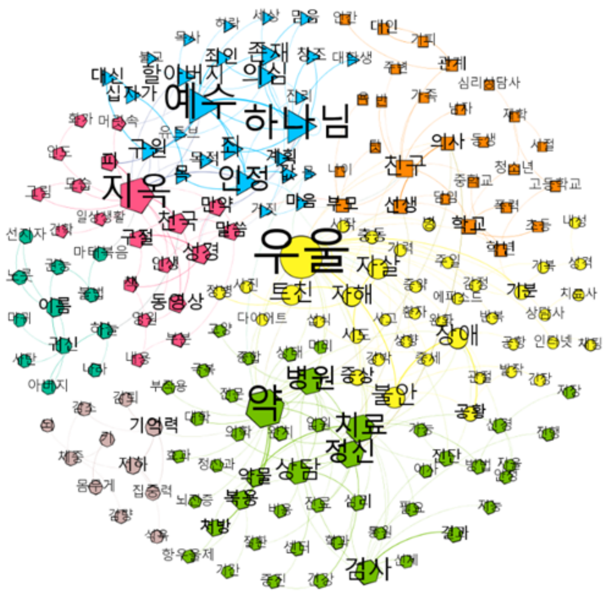
<Table 8> Community detection results of non-depression related texts

Community 0	Community 1	Community 2
Treatment, Medication, Patient, Behavior, Help, Effect, Method, Need, Cognitive, Improvement, Medicinal, Use, Taking, Overcoming, Important, Cancer, Progress, Possible, Therapy, Law, Prescription, Parallel, Herbal, Start, Active, Surgery, Excess, Alcohol, Acupuncture, Expectation, Role, Remind, Attention, Period, Illness, Diet	Symptoms, Disorders, Depression, Anxiety, Mood, Emotions, State, Causes, Changes, Diagnosis, Bipolar, Persistent, Major, Panic, Occurrences, Accompanied, Postpartum, Relief, Negatives, Experiences, Insomnia, Tick, Control, Criteria, Coaching, Prenatal, deterioration, expression, improvement, personality, feeding, accuracy, appeal, personality, psychosis, fundamental, anxiety, transition, ups and downs, representation, grasp	Decrease, weight, decrease, increase, appetite, risk, memory, concentration, motivation, loss, interest, ability, improve, loss, pleasure, decline, sluggish
Community 3	**Community 4**	**Community 5**
Suicide, Prevention, School, Violence, Attempt, Insurance, Dementia, Death, Impulse, Addiction, Accident, Elementary, Korea, Disaster, Death, Game, Prayer, Paid, Repetition	Mental, health, disease, body, brain, medicine, doctor, activity, insurance, maintenance, immunity, science, hospital, management, checkup, influence, heart, department, promotion, decision, system, cardiovascular, modern, schizophrenia, physical, Rescue, people, service, food, pain	Counseling, psychology, center, factor, welfare, support, expert, inspection, deletion, environment, elderly, biology

depressive symptoms that included words such as “depressed,” “weight,” and “increased”.

The simultaneous occurrence word network of the latent depression literature class includes 184 nodes and 310 edges, and consists of a total of 7 communities. The results of visualizing the network of simultaneous occurrences of the latent depression literature are shown in <Figure 4> below, and words belonging to each community are shown in <Table 9>. It was ‘depressed’ with the greatest linking centrality in the co-occurrence word network, followed by ‘medicine,’ ‘hell,’ and ‘God’. Unlike the non-depressing literature class, words related to religion ranked high in the center of connection. In addition, community 1 that includes ‘pastor,’ community 4 that includes words such as ‘Matthew,’ and community 5 that includes ‘phrase’ and ‘the Bible’ included religious words. Community 0, which includes ‘depressed’ and ‘counselor,’ is composed of words related to depression symptoms and treatment, and community 3, which includes ‘medicine’ and ‘treatment,’ is composed of words related to treatment of depression. Community 2 includes words related to family and interpersonal relationships, such as ‘friend,’ ‘around,’ and ‘family’.

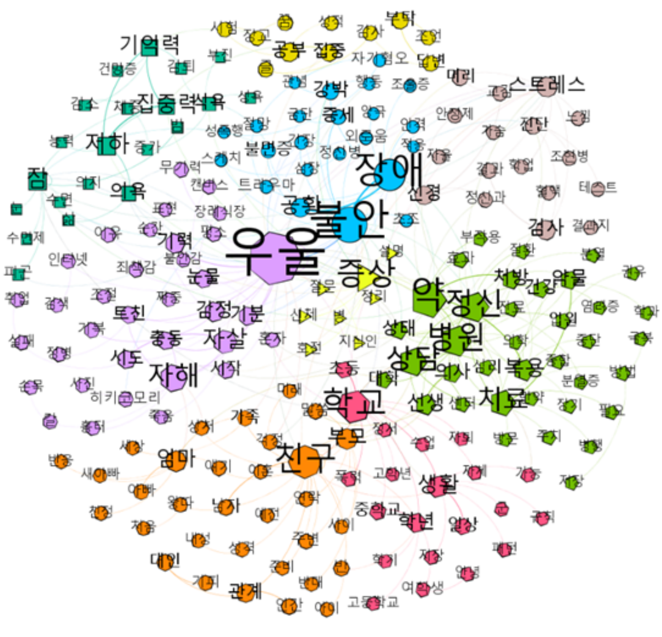
The simultaneous occurrence word network of the depression tendency literature class is composed of 205 nodes and 352 edges, and there are a total of 9 communities. The result of visualizing the network of simultaneous occurrences of the corresponding class is shown in <Figure 5>, and the words belonging to each community are shown in <Table 10> below. The word with the greatest connection centrality in the network was ‘depression,’ followed by ‘symptoms,’ ‘anxiety,’ and ‘hospitals’. In addition, as words such as ‘friend’



〈Figure 4〉 Co-word network in latent depression related texts

〈Table 9〉 Community detection results of latent depression related texts

Community 0	Community 1	Community 2	Community 3
Melancholy, trauma, anxiety, disability, symptom, panic, attempt, suicide, self-harm, mood, impulse, energy, affection, patient, photograph, symptoms, symptoms, symptoms, feelings, tension, internet, week, ups and downs, counselor, therapist, sick, Diet, start, obsessive, repetitive, propensity, episode, eating, joint, personality, accident, chat, airport, tolerance, seizure, relief	Jesus, God, Doubt, Being, Salvation, Sin, Instead, Cross, Sinner, Recognition, Value, Neck, Grandfather, Creation, Plan, Purpose, College Student, Faith, Pastor, World, Heart, False, Youtub, Truth, Permission, Buddhism, body	Teacher, doctor, friend, elementary, school, avoidance, interpersonal, violence, relationship, grade, class, human, man, middle school, parent, family, high school, attending, psychological counselor, youth, swear, brother, age, homeroom teacher, peripheral, Blame, times	Medicine, Taking, Prescription, Medication, Treatment, Examination, Results, Diagnosis, Hospital, Counseling, Mental, Psychological, Health, College, Clinic, Medicine, Heal, Center, Method, Overcoming, Body, Neurological, Condition, Hospitalization, Side Effects, Psychiatry, antidepressant, effect, cost, need, stability, promotion, uplift, outpatient, disease, epilepsy, progression, professional, head, autonomy, possible, storage, department, period, intelligence, comprehensive, director
Community 4	Community 5	Community 6	
Ghost, name, father, heaven, role, prophet, power, lawlessness, matthew, devil, satan, kingdom	Verse, Bible, Blood, Hell, Heaven, If, Video, Book, Life, Picture, Look, Word, Daily Life, Head, Eternity, Field Trip, India, Painter, Content, Part	Memory, decline, decline, weight, height, concentration, appetite, weight, loss, decrease, brain	



〈Figure 5〉 Co-word network in high depression related texts

<Table 10> Community detection results of high depression related texts

Community 0	Community 1	Community 2	Community 3	Community 4
Anxiety, disability, panic, obsessive, symptoms, insomnia, tension, trauma, loneliness, despair, psychosis, self-loathing, behavior, agitation, personality, sketch, bipolar, manic depression, withdrawal, sexual assault, notion, adaptation, heart	Melancholy, self-harm, suicide, energy, emotion, mood, attempt, impulse, tears, tramp, start, ups and downs, hikikomori, helplessness, alone, reason, internet, guilt, knife, usual, control, failure, scar, spirit, irritation, Photo, canvas, search, funeral home, death, moment, employment, anxiety, expression, wrist	Decline, sleep, concentration, memory, motivation, appetite, tiredness, decline, life, decrease, night, weight, increase, sleep, will, sleeping pills, eyes, forgetfulness, libido, sluggishness, ability	About, Hospital, Mental, Counseling, Treatment, Taking, Teacher, Medication, Doctor, Prescription, State, Health, College, Hospitalization, Psychological, Effect, Center, Method, Clinic, Disease, Side Effect, Need, Recommend, Overcome, Medicine, Synthesis, disruption, visit, interruption, department, schizophrenia, attending physician, storage, if, organ, parallel, anxiety	Friend, Parent, Mom, Relationship, Interpersonal, Family, Man, Human, Word, Half, Worry, Dad, Avoid, Contact, Talk, Around, Personality, Bullying, Hurt, Relationship, Child, Parent, Old, Opposition, Stepdad, Introspective, divorced, first, ready, future, reaction, world
Community 5	**Community 6**	**Community 7**	**Community 8**	
Symptom, body, disease, question, cleanup, improvement, explanation, intellectual	School, Life, Grade, Elementary, Daily, Middle School, Violence, Schoolgirl, Possible, Goodbye, Disorder, Withdrawal, Self, Class, Semester, Pattern, Military, High School, Emotion, Rule, Seniors	Study, concentration, request, answer, gratitude, exam, dream, writing, officer, advice, grades	Stress, Examination, Neuron, Diagnosis, Head, Results, Sympathetic, Psychiatry, Academic, Stabilizer, Autonomy, Schizophrenia, Results sheet, Blood, Feeling, Test, Chest	

and 'hospital' are located above the connection center, it can be seen that words related to symptoms and treatment of depression are important in the depressive tendency literature class. Community 1 of the depressive tendency literature was composed of words related to suicidal thoughts, attempts, and depressive symptoms, and Community 0 and Community 2 were composed of words related to the DSM-5 major depressive disorder diagnosis criteria. In addition, in communities 4 and 6, words related to family and social relationships such as 'parents,' 'persons,' 'relationships,' and 'schools' were identified. This is one of the things that can persist or cause depression. In addition, there were communities 3 and 8 consisting of "hospital," "counseling," and "test" related to the treatment of depression, and communities 5 and 7 were composed of words asking questions about the symptoms of depression of oneself or others.

In summary, the non-depressive literature class was composed of contents related to the treatment of depression in general, and the latent depressive literature class included contents related to religion in addition to symptoms and treatment of depression. In the case of the latent depressive literature class with relatively mild depression symptoms, it is assumed that religion-related content predominates over the depressive tendency literature because it can relieve depression by relying on religion. The depressive tendency literature was composed of contents related to depressive symptoms in general, and in addition, questions about the cause, treatment, and symptoms of depression appeared.

4.3 Depression tendency classification

In this study, previously annotated corpuses were used as experimental data, 80% of the corpus were used for training, 20% were used as test data for performance evaluation, and data were separated through random extraction. In all experiments, the input length was set 256 words, the epoch was set to 10, the batch size was set to 32, and the optimizer was set to Adam. The glove model used in this study used the learned glove model by using the corpus morphologically analyzed by the Korean Wikipedia, KorQuAD, and Naver movie corpus with Mecab (Lee, 2019). In addition, in the case of FastText, the Korean model released on Facebook was used. <Table 11> shows the performance of using only the text embedded in the experiment setting described above in the depression tendency classification model and the performance when the model was trained by combining various text features together. Accuracy was used as a measure for performance evaluation.

As a result, the performance when learning the depression tendency classification model using only text in common was the lowest, and among them, the lowest accuracy rate of 77.52% was shown when using the Bi-LSTM algorithm and the FastText word embedding model. However, as the model was trained by combining text features, it gradually improved its performance. When the text and sentiment scores were combined, all algorithms showed a performance improvement of 0.32% to 1.2%, and when using the

〈Table 11〉 Results of a depression tendency classification model in accuracy rates

Additional Feature \ Deep-Learning Algorithm And Embedding	CNN + Glove	CNN + FastText	Bi-LSTM + Glove	Bi-LSTM + FastText	Multilin-gual BERT	KorBERT
Text	78.28%	80.02%	78.07%	77.52%	79.69%	<u>81.22%</u>
Text + Sentiment Score	79.48%	80.78%	78.39%	77.96%	80.35%	<u>82.19%</u>
Text + Topic	78.93%	80.67%	78.39%	76.22%	80.13%	<u>82.08%</u>
Text + Sentiment Score + Topic	81.11%	81.98%	79.37%	81.22%	80.89%	<u>83.28%</u>

CNN algorithm and the Glove word embedding model, the biggest performance improvement was 1.2%. When the model was trained by combining text and topics, most of the algorithms showed improved performance, but the Bi-LSTM algorithm and FastText word embedding model decreased by 1.3%. In combining text and topics, KorBERT using MeCab POS Tagger showed the greatest performance improvement of 0.86%. When the model was trained by combining all of the text, sentiment scores, and topics, the performance of all algorithms improved, showing a large improvement range of 1.2% to 3.7%. Among them, the performance of the Bi-LSTM algorithm and FastText word embedding model showed the greatest performance improvement of 3.7%. Therefore, it can be seen that the deep learning-based depression tendency classification model combining text and text features outperforms the existing text-based classification model.

There was a maximum difference in performance of 1.74% between the sentiment score used as the text feature and the topic, and it was judged that the sentiment score rather than the topic helped improve the performance. In the case of using both the sentiment score and the topic as text features, performance was improved compared to the case of using only the sentiment score and topic, respectively. There was a performance improvement of at least 0.54% and up to 5%, and among them, the Bi-LSTM algorithm and FastText word embedding showed the greatest improvement of 5%. Therefore, it can be seen that the use of both emotional scores and topics as text features helped improve performance.

Overall, the model with the highest accuracy rate was KorBERT, and the highest accuracy rate, 83.28%, was achieved when depressive tendencies were classified by combining emotion scores and topics with text. After KorBERT, the CNN algorithm and FastText word embedding model showed an accuracy rate of 81.98%, which is 1.3% lower than KorBERT. Taken together, it can be seen that learning in the KorBERT algorithm by combining all of the emotion scores and topics together with the text can achieve the best performance in classifying depression trends.

5. Conclusion

This study attempted to classify depression trends using Korean social media texts, which were not performed much in previous studies related to automatic depression detection. At this time, the performance was improved by performing deep learning-based text classes using various text features such as embedded text, sentiment score, and topic. In addition, by analyzing corpuses that have been verified by a psychiatrist for each class, we tried to understand the characteristics of each class and the differences between classes.

As a result of classifying depression trends, it was confirmed that the highest accuracy rate, 83.28%, was obtained when texts were embedded using KorBERT and combined with emotional scores and topics to classify depression trends in Korean social media texts. In addition, regardless of the type of deep learning algorithm, when the depression tendency classification was performed using only the embedded text, the lowest accuracy was shown. However, as the emotion score and topic were combined with the embedded text and learned, the performance gradually improved. Through this, it was confirmed that the use of various text features improves the accuracy rate of classification of depression tendency. We provided a basis for more accurate depressive tendency classification by tuning the classification model using text features to better performance. The good performance enables rapid and accurate classification of depressive tendencies, enabling faster detection of depressive tendencies prevalent in Korean society.

The comprehensive implications of the study are as follows. First, this study contributed to the detection of depression in Korean social media through Korean-based depression detection. This is significant in that automatic depression detection studies were able to improve the limitations that were mainly based on English. Second, previous studies (Alessa, Faezipour, & Alhassan, 2018; Cheng & Chen, 2019; Al Essa, 2018; Lilleberg, Zhu, & Zhang, 2015; Pasupa & Ayutthaya, 2019) found that using various text features for classification improves performance. We were able to confirm the results reported in these previous studies. In this study, it was confirmed that there is an improvement in accuracy of 1.2% to 3.7% when all the various text features are combined and classified compared to when only embedded text is used for classification.

The limitations of this study are as follows. Among the data sources, tweeters composed of short texts have a limit in classification because they are composed of words less than 256 words, which is a fixed input length, and fill the remaining length with zeros. Therefore, in subsequent studies, if the model is trained according to the length and characteristics of the data source, more significant results can be expected. In addition, in addition to the text features used in this study, it is expected that performance improvement can be expected if additional text features such as sentiment score, topic effective in classifying depression tendencies are used.

References

Aizawa, A. (2003). An information-theoretic perspective of tf–idf measures. Information Processing & Management, 39(1), 45-65. http://doi.org/10.1109/ICHI.2018.00058

Al Essa, A. (2018). Efficient Text Classification with Linear Regression Using a Combination of Predictors for Flu Outbreak Detection. Doctoral dissertation, University of Bridgeport.

Alessa, A., Faezipour, M., & Alhassan, Z. (2018). Text classification of flu-related tweets using fasttext with sentiment and keyword features. In 2018 Institute of Electrical and Electronics Engineers International Conference on Healthcare Informatics (ICHI), 366-367. http://doi.org/10.1109/ICHI.2018.00058

Athiwaratkun, B., Wilson, A. G., & Anandkumar, A. (2018). Probabilistic fasttext for multi-sense word embeddings. arXiv. https://doi.org/10.48550/arXiv.1806.02901

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. The Journal of Machine Learning Research, 3, 993-1022. https://doi.org/10.1016/b978-0-12-411519-4.00006-9

Blondel, V. D., Guillaume, J. L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. Journal of Statistical Mechanics: Theory and Experiment, 2008(10), P10008. https://doi.org/10.1088/1742-5468/2008/10/P10008

Callon, M., Courtial, J. P., Turner, W. A., & Bauin, S. (1983). From translations to problematic networks: An introduction to co-word analysis. Social Science Information, 22(2), 191-235. https://doi/org/10.1177/053901883022002003

Cheng, C. H. & Chen, H. H. (2019). Sentimental text mining based on an additional features method for text classification. PloS One, 14(6), e0217591. https://doi.org/10.1371/journal.pone.0217591

Chronis, G. & Erk, K. (2020). When is a bishop not like a rook? When it's like a rabbi! Multi-prototype BERT embeddings for estimating semantic relationships. In Proceedings of the 24th Conference on Computational Natural Language Learning, 227-244. https://doi.org/10.18653/v1/2020.conll-1.17

Conway, M. & O'Connor, D. (2016). Social media, big data, and mental health: current advances and ethical implications. Current Opinion in Psychology, 9, 77-82. https://doi.org/10.1016/j.copsyc.2016.01.004

Coppersmith, G., Dredze, M., & Harman, C. (2014, June). Quantifying mental health signals in Twitter. In Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality, 51-60. https://doi.org/10.3115/v1/w14-3207

Coppersmith, G., Dredze, M., Harman, C., Hollingshead, K., & Mitchell, M. (2015). CLPsych 2015 shared task: Depression and PTSD on Twitter. In Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality, 31-39. https://doi.org/10.3115/v1/w15-1204

De Choudhury, M., Kiciman, E., Dredze, M., Coppersmith, G., & Kumar, M. (2016). Discovering shifts to suicidal ideation from mental health content in social media. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, 2098-2110. https://doi.org/10.1145/2858036.2858207

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv. http://arxiv.org/abs/1810.04805

Friedrich, M. J. (2017). Depression is the leading cause of disability around the world. Jama, 317(15), 1517-1517. https://doi.org/10.1001/jama.2017.3826

Guntuku, S. C., Yaden, D. B., Kern, M. L., Ungar, L. H., & Eichstaedt, J. C. (2017). Detecting depression and mental illness on social media: an integrative review. Current Opinion in Behavioral Sciences, 18, 43-49. https://doi.org/10.1016/j.cobeha.2017.07.005

Kim Y. (2014). Convolutional neural networks for sentence classification. EMNLP2014-2014 Conference on Empirical Methods in Natural Language Processig, Association for Computational Linguistics, 1746-1751. https://doi.org/10.3115/v1/d14-1181

KNU Korean Emotion Dictionary (2018, November 5). Available: https://github.com/park1200656/KnuSentiLex

Lalithamani, N., Thati, L. S., & Adhikesavan, R. (2014). Sentence level sentiment polarity calculation for customer reviews by considering complex sentential structures. IJRET: International Journal of Research in Engineering and Technology, 3(3), 433-438. https://doi.org/10.15623/ijret.2014.0303081

Lee G. (2019). Korean Ebedding. Korea: Acorn Publishing.

Liang, H., Sun, X., Sun, Y., & Gao, Y. (2017). Text feature extraction based on deep learning: a review. EURASIP Journal on Wireless Communications and Networking, 2017(1), 1-12. https://doi.org/10.1186/s13638-017-0993-1

Lilleberg, J., Zhu, Y., & Zhang, Y. (2015). Support vector machines and word2vec for text classification with semantic features. In 2015 IEEE 14th International Conference on Cognitive Informatics & Cognitive Computing (ICCI* CC), 136-140. Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/ICCI-CC.2015.7259377

Lim, J. H., Kim, H. K., & Kim, Y. K. (2020). Recent R&D trends for pretrained language model. Electronics and Telecommunications Trends, 35(3), 9-19. https://doi.org/10.22648/ETRI.2020.J.350302

Liu, G. & Guo, J. (2019). Bidirectional LSTM with attention mechanism and convolutional layer for text classification. Neurocomputing, 337, 325-338. https://doi.org/10.1016/j.neucom.2019.01.078

Martin, L., Muller, B., Suárez, P. J. O., Dupont, Y., Romary, L., de la Clergerie, É. V., Seddah, D., & Sagot, B. (2019). Camembert: a tasty french language model. https://doi.org/10.18653/v1/2020.acl-main.645

Moon, E. & Han, S. (2011). A qualitative method to find influencers using similarity-based approach in the blogosphere. International Journal of Social Computing and Cyber-Physical Systems, 1(1), 56-78. https://doi.org/10.1504/ijsccps.2011.043604

Mowery, D., Smith, H., Cheney, T., Stoddard, G., Coppersmith, G., Bryan, C., & Conway, M. (2017). Understanding depressive symptoms and psychosocial stressors on Twitter: a corpus-based study. Journal of Medical Internet Research, 19(2), e48. https://doi.org/10.2196/jmir.6895

Nam, K. K., Ackerman, M. S., & Adamic, L. A. (2009). Questions in, knowledge in? A study of Naver's question answering community. In Proceedings of the SIGCHI Conference on Human Factors in

Computing Systems, 779-788. https://doi.org/10.1145/1518701.1518821

Orabi, A. H., Buddhitha, P., Orabi, M. H., & Inkpen, D. (2018). Deep learning for depression detection of twitter users. In Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic, 88-97. https://doi.org/10.18653/v1/W18-0609

Pasupa, K. & Ayutthaya, T. S. N. (2019). Thai sentiment analysis with deep learning techniques: A comparative study based on word embedding, POS-tag, and sentic features. Sustainable Cities and Society, 50, 101615. https://doi.org/10.1016/j.scs.2019.101615

Pennington, J., Socher, R., & Manning, C. D. (2014). Glove: Global vectors for word representation. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), 1532-1543. https://doi.org/10.3115/v1/D14-1162

Petterson, J., Smola, A. J., Caetano, T. S., Buntine, W. L., & Narayanamurthy, S. M. (2010). Word features for latent dirichlet allocation. In NIPS, 1921-1929. https://doi.org/10.1.1.942.7045

Qaiser, S. & Ali, R. (2018). Text mining: use of TF-IDF to examine the relevance of words to documents. International Journal of Computer Applications, 181(1), 25-29. https://doi.org/10.5120/ijca2018917395

Resnik, P., Armstrong, W., Claudino, L., Nguyen, T., Nguyen, V. A., & Boyd-Graber, J. (2015). Beyond LDA: exploring supervised topic modeling for depression-related language in Twitter. In Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality, 99-107. https://doi.org/10.3115/v1/w15-1212

Resnik, P., Garron, A., & Resnik, R. (2013). Using topic modeling to improve prediction of neuroticism and depression in college students. In Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing, 1348-1353. url: https://www.aclweb.org/anthology/D13-1133

Ruas, T., Ferreira, C. H. P., Grosky, W., de França, F. O., & de Medeiros, D. M. R. (2020). Enhanced word embeddings using multi-semantic representation through lexical chains. Information Sciences, 532, 16-32. https://doi.org/10.1016/j.ins.2020.04.048

Schwartz, H. A., Eichstaedt, J., Kern, M., Park, G., Sap, M., Stillwell, D., Kosinski, M., & Ungar, L. (2014). Towards assessing changes in degree of depression through facebook. In Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality, 118-125. https://doi.org/10.3115/v1/w14-3214

Tadesse, M. M., Lin, H., Xu, B., & Yang, L. (2019). Detection of depression-related posts in reddit social media forum. IEEE(Institute of Electrical and Electronics Engineers) Access, 7, 44883-44893. https://doi.org/10.1109/ACCESS.2019.2909180

Trotzek, M., Koitka, S., & Friedrich, C. M. (2018). Early detection of depression based on linguistic metadata augmented classifiers revisited. In International Conference of the Cross-Language Evaluation Forum for European Languages, 191-202. Springer, Cham. https://doi.org/10.1007/978-3-319-98932-7_18

Tsugawa, S., Kikuchi, Y., Kishino, F., Nakajima, K., Itoh, Y., & Ohsaki, H. (2015). Recognizing depression from twitter activity. In Proceedings of the 33rd annual ACM conference on Human Factors in Computing Systems, 3187-3196. https://doi.org/10.1145/2702123.2702280

Turney, P. D. & Pantel, P. (2010). From frequency to meaning: Vector space models of semantics. Journal of Artificial Intelligence Research, 37, 141-188. https://doi.org/10.1613/jair.2934

Wang, Z. Y., Li, G., Li, C. Y., & Li, A. (2012). Research on the semantic-based co-word analysis. Scientometrics, 90(3), 855-875. https://doi.org/10.1007/s11192-011-0563-y

World Health Organization (2020). Available: https://www.who.int/health-topics/depression#tab=tab_1

Yin, Z. & Shen, Y. (2018). On the dimensionality of word embedding. arXiv preprint arXiv:1812.04224. https://doi.org/10.48550/arXiv.1812.04224

Yun-tao, Z., Ling, G., & Yong-cheng, W. (2005). An improved TF-IDF approach for text classification. Journal of Zhejiang University-Science A, 6(1), 49-55. https://doi.org/10.1007/BF02842477

Zhang, L., Huang, X., Liu, T., Li, A., Chen, Z., & Zhu, T. (2014). Using linguistic features to estimate suicide probability of Chinese microblog users. In International Conference on Human Centered Computing, 549-559. Springer, Cham. https://doi.org/10.1007/978-3-319-15554-8_45

Zhao, J., Zhou, Y., Li, Z., Wang, W., & Chang, K. W. (2018). Learning gender-neutral word embeddings. arXiv preprint arXiv:1809.01496. https://doi.org/10.18653/v1/d18-1521