

SENTIMENT ANALYSIS OF DEPRESSION DISCOURSE: A CORPUS-BASED STUDY OF NOTES FROM REDDIT

Shehzaib Yasin¹

Department of English, University of Sialkot, Sialkot, Province of Punjab, Pakistan

Email: yshehzaib@gmail.com¹

Shiza Shakeel²

Department of English, University of Sialkot, Sialkot, Province of Punjab, Pakistan

Email: shizanoor420@gmail.com²

Muhammad Sabboor Hussain³

Professor, Department of English, University of Sialkot, Punjab, Pakistan.

Email: dr.muhamadsabboorhussain@uskt.edu.pk

ABSTRACT

This study aims to explore the language used by the depressed community and its effects on readers' sentiments. This study examines the DNC (Depression Notes Corpus), which was scraped using Python. This corpus consisted of depression notes scraped from the social media platform Reddit between 2013 and 2023 (100 posts per year). This platform was chosen to avoid bias and maintain the anonymity of the population and sample. The purpose of this study is to conduct a detailed corpus-based sentiment analysis of notes on depression. The research question for this study is: Which lexical categories do depressed people use most, and what are the sentiments of depressed people when they write (positive, negative or neutral)? The objectives of this study are to explore the use of lexical categories and to study the impact of choices made by depressed people on others' sentiments. The depression notes were analyzed by using the corpus tool Sketch Engine. In the findings, the study revealed that the most frequent lexical category in the corpus was the verb, which suggests that depressed people want to express action. This was analyzed by examining the notes of Redditors. They mostly use adjectives to beautify their writing and to create an impact on others' sentiments. At termination, the study concludes that all positive words have positive effects on others' sentiments, negative words have negative effects, and neutral words have no effect.

Keywords: *sentiment analysis, depression discourse, depression notes, Reddit*

1. INTRODUCTION

1.1 Background of the Study:

Depression has become a more common but serious health condition in today's materialistic world. Almost all the people (of all ages) are depressed because of their issues and complexities of this very life. People may feel tired, hopeless and worthless. Some are depressed for jobs, bread, houses, and calm in their lives, while some are materially wealthy but also depressed for being happy. Happiness is now becoming rare in our society. Despite being depressed, depressed people need someone to purify their emotions (catharsis). Unfortunately, nowadays, finding a loyal listener is also a challenge. So these people keep on thinking about their situations and become weaker mentally, physically and emotionally. For their catharsis, they use some online resources to let the depression go away. Social media has become a popular and significant tool for communication today. People share millions of messages, thoughts, issues, feelings, and emotions on social media worldwide in the form of posts, pictures, and comments to purify their emotions. The most popular social media platforms are Facebook, Tumbler, Instagram, Twitter and Reddit etc. The composers of those messages write about their lives, opinions, and others' lives as well. They depict their sadness, loneliness, hopelessness, worthlessness and valueless existence in society. These platforms allow them to express their feelings and emotions freely.

Some platforms may become dangerous to use, and some become user-friendly. Marginalized depressed people express themselves more powerfully so that the researchers may find increased sadness in them after researching the other people online who express themselves strongly online. A systematic collection of data, whether written or spoken, is called a corpus. Corpus linguistics is the branch of linguistics which deals with the study and analysis of large amounts of data by using different corpus tools. Sentiment refers to the emotional tone or attitude expressed in a piece of text, speech, or communication. It can be categorized as positive, negative, or neutral, reflecting the author's subjective feelings, opinions, or reactions toward a particular subject or topic. Sentiment analysis aims to determine and quantify sentiment in given content automatically. Sentiment analysis is a natural language processing technique used to determine the sentiment expressed in a piece of text, such as positive, negative, or neutral. It involves analyzing and interpreting the subjective information present in the text to understand the overall sentiment or opinion conveyed by the writer or speaker.

1.2 Scope and Significance of Study:

This study is of great scope as it aims to explore a marginalized community, which is a depressed community, and aims to transform society by talking or treating them gently. This study is very significant because of its unique features, such as extracting data anonymously and analyzing the emotions and sentiments without manipulating them. This study will contribute solely to the transformation of society.

1.3 Statement of the Problem:

In recent years with technological advancements, depressed people have also become marginalized. They live alone mostly and struggle for their mental peace. This study tackles the problem of depressed people, explaining why depressed people are marginalized and why they are not treated well. What type of language do they use when they are depressed, and what are their intentions?

1.4 Aim and Objective of the Study:

This study aims to explore the use of language by depressed people and its impact on the sentiments of others. The objectives of this study are:

1. To explore the usage of lexical categories by depressed people.
2. To study the impact of the choices made by depressed people on the sentiments of others.

1.5 Research Questions:

This study aims to answer the following research questions:

1. Which lexical categories do the depressed people use the most?
2. What are the sentiments of depressed people when they write? (Positive, negative or neutral).

1.6 Delimitations and Limitations of the Study:

Due to time constraints, this study is limited to the online platform Reddit for data collection. This study is limited to posts from the depressed community on Reddit from 2013 to 2023 (100 posts/year). This study is also limited to the sentiment analysis of these depression notes.

2. LITERATURE REVIEW

2.1 Recent Studies:

Recent studies show that corpus-based research has emerged in this evolving landscape. Most researchers began conducting corpus-based studies using technology. Studies from 2010 to 2025 show that depression has become the most considerable sensitive issue which should be studied and dealt with great care. People have also been working on sentiment analysis across various fields. The following are the findings from the recent studies.

With the proliferation of blogs and social networks, sentiment analysis became a field of interest for many researchers. A very broad overview of the existing work was presented in Pang & Lee (2008). In their survey, the authors describe existing techniques and approaches for opinion-oriented information retrieval. However, few studies in opinion mining have considered blogs, and even fewer have addressed microblogging. In Yang et al. (2007), the authors use weblogs to construct a corpus for sentiment analysis and use emotion icons assigned to blog posts as indicators of users' moods. The authors applied SVM and CRF learners to classify sentiment at the sentence level and then investigated several strategies for determining the overall sentiment of the document. As a result, the winning strategy is to treat the sentiment of the document's last sentence as the document-level sentiment.

J. Read (Read, 2005) used emoticons such as “:-)” and “:- (“ to form a training set for the sentiment classification. For this purpose, the author collected texts containing emoticons from Usenet newsgroups. The dataset was divided into “positive” (texts with happy emoticons) and “negative” (texts with sad or angry emoticons) samples. Emoticons-trained classifiers: SVM and Naive Bayes were able to achieve up to 70% accuracy on the test set. In Go et al. (2009), the authors used Twitter to collect training data and then perform a sentiment search. The approach is similar to that of Read (2005). The authors construct corpora by using emoticons to obtain “positive” and “negative” samples, and then use various classifiers. The best result was obtained using the Naive Bayes classifier with mutual information for feature selection. The authors achieved up to 81% accuracy on their test set. However, the method showed poor performance across three classes (“negative”, “positive” and “neutral”).

In this regard, the social media platform offers both benefits and obstacles. Positively, it provides the writer with anonymity, enabling them to express their emotions freely (Rajadesingan et al., 2015). In recent years, sentiment analysis has gained favour among research communities. Opinion analysis and opinion mining are other names for sentiment analysis. The sentiment analysis task has increased recently. A summary of opinion mining has already been provided by various sentiment analysis research efforts (Lighthart et al., 2021). The study topic from 2000 to 2015 is covered in Piryani et al. (2017), which offers a framework for computationally processing unstructured data, with the main objective of collecting opinions and determining their moods.

As of this writing, there are two distinct implementations of our general approach. In both cases, the analyst specifies two sets of “seed words,” one positive and one negative. Next, we use our small seed sets of positive and negative phrases, along with the actual opinion texts, to generate dictionaries using two methods: word conjunctions and word co-occurrence. The concepts put out by McIntyre and Walker (2019) to perform a thorough corpus-based analysis of a collection of suicide notes that have been obtained from Reddit have an impact on this work. Leech and Short's (2007) checklist was used to determine style categories for the stylistic analysis in this study, which employs both quantitative and qualitative methodologies to examine the linguistic aspects of the data.

2.2 Research Gaps:

A review of the related literature found that no research had examined depression discourse beyond clinical research. Those researchers who were found selected posts from other social media platforms, such as Facebook, Tumblr, and Instagram, rather than from Reddit.

2.2.1 Methodological Gaps:

In the studies reviewed, many methodological gaps were found. No proper methodology was used in the reviewed articles, and no methodology section was present in the studies.

2.2.2 Conceptual Gaps:

In recent studies, a lack of conceptual frameworks was found. It seemed that researchers were not aware of what they were doing.

2.3 Theoretical Framework:

In the recently viewed studies, a theoretical framework was lacking in most. In this study, the Appraisal Theory of emotion was used to guide the data analysis process.

2.4 Conceptual Framework:

Recent studies lack a conceptual framework, which is indeed a gap in the study. For this research, a conceptual framework for extracting lexical categories and adjectives was used.

2.5 Analytical Framework:

In the analytical framework, many mathematical models were identified, and the data were analysed by examining their mathematical values. In this study, an analytical framework based on Sketch Engine was used to guide the research.

3. METHODOLOGY

Methodology concerns the methods used to conduct a study. It consists of the following parts:

3.1 Research Design:

For this study, a convergent mixed methods research design was used for data collection and analysis. Both qualitative and quantitative data were collected and analysed simultaneously.

3.2 Research Tools (Instruments and Items)

Throughout this study, depression notes serve as the research tools and instruments. These tools or items were used to conduct this research.

3.3 Population and Sampling:

A population is the entire community that a researcher wants to study. For the entire study, the population is a depressed community. A sample is a representative of the population, and in this study, it was randomly selected from the Reddit online platform.

3.4 Data Collection Procedure

The Reddit API was used to scrape data from depression notes from 2013 to 2023 (100 posts per year) using Python on a computer. Reddit was chosen for data collection because it is a genuine source, and people's identities remain hidden, allowing them to use it anonymously. Reddit has a feature that lets us post and comment anonymously, without revealing our identities, and most people share the truth there without fear. In this way, we obtained genuine data from people without revealing their identities.

4. DATA ANALYSIS

Firstly, the online tool Sketch Engine was used to determine the total frequency of the word in the Depression Notes Corpus (DNC). Then the analysis of the lexical categories of the words found in the DNC, along with their frequencies. The following table #2 shows the total frequency of words along with the total number of tokens in DNC.

Name	Numbers
Total Number of Tokens	2,280
Frequency of all the Tokens	1,85,420

Table #2 showing the frequencies of tokens

4.1 Training the classifier

As researchers wanted to analyse the corpus with the help of a computer, prior information on how to analyse the data is given to the computer software or website to run it smoothly, which is known as Training the Classifier. It takes several steps to train a computer to use it as a classifier. Those steps are mentioned below:

4.2. Feature extraction

The collected dataset is used to extract features for training our sentiment classifier. In this study, we will extract the following features to support our arguments.

For 1st research question, the following features were extracted.

Lexical Categories: Lexical categories are the basic categories of words which are used according to their functions. There are eight types of lexical categories, which are mentioned below:

Noun: The name of a person, place, thing or idea is called a noun. E.g. boy, book, city, etc.

Pronoun: A word which is used instead of a noun is called a pronoun.

E.g. I, me, he. She, it, etc.

An adjective: A word which qualifies a noun or pronoun is called an adjective—E.g, beautiful, ugly, etc.

Verb: A word which shows an action or connects the subject with another part of the sentence is called a verb.

E.g. go, run, is, am, are, etc.

Adverb: A word which qualifies a verb, an adjective or another adverb is called an adverb. E.g. beautifully, smoothly, garden, etc.

Conjunction: A word which connects words, phrases or clauses is called a conjunction. E.g. and, or, for, etc.

Interjection: A word used to express sudden emotion is called an interjection. E.g. hurrah!, alas! Etc.

Preposition: A preposition is a word or group of words used before a noun, pronoun, or noun phrase to show direction, time, place, location, spatial relationships, or to introduce an object E.g. in, on, under, etc.

For the 2nd question, we extracted an adjective feature to determine the sentiments of depressed people.

Adjectives are words that qualify nouns; they can also be used to analyse people's sentiments through different uses.

Positive Sentiment: A positive sentiment is a favourable attitude or emotion which is expressed towards any topic or something. E.g. good, successful, etc.

Negative Sentiment: A negative sentiment is an unfavourable attitude or emotion which is expressed towards any topic or something. E.g. bad, ugly, etc.

Neutral Sentiment: A neutral sentiment is neither a favourable nor unfavourable attitude nor emotion which is expressed towards any topic or something. E.g. same, constant, etc.

5. FINDINGS AND DISCUSSIONS:

The following findings were observed:

5.1 Lexical Category Analysis:

Lexical categories are categories of words that are used according to their function. These categories include noun, pronoun, verb, adverb, conjunction, interjection and interjections. We

collected the corpus from the Reddit API to analyze the notes of depression which depressed people write. The DNC data consists of 1, 85,420 words/tokens and 2,280 distinct token types. Some findings related to lexical category analysis are described below:

1. In DNC, 794 types of tokens were found for Nouns, with a frequency of 27,679 tokens.
 2. In the DNC, 23 types of tokens were found for pronouns, with a frequency of 22,530.
 3. In DNC, 360 types of Adjective tokens and 10,574 token frequency.
 4. In the DNC, 480 types of verbs were found, while 38,576 tokens of frequency were found.
 5. In the DNC, 212 types of Adverb tokens were found, with a frequency of 17,283.
 6. In the DNC, 7 types of Conjunction tokens were found, with a frequency of 6,682.
 7. In the DNC, 10 types of Number tokens were found, with a frequency of 2,050.
 8. In the DNC, 55 types of Preposition tokens were found, with a frequency of 16,042.
 9. In DNC, 339 types of others' tokens were found, with a frequency of 44,044 tokens.
- (Other category includes all the punctuations, contractions and incomplete words.)

Chart 5.1 a shows the frequency of the tokens

Chart 5.1 b shows the number of token types

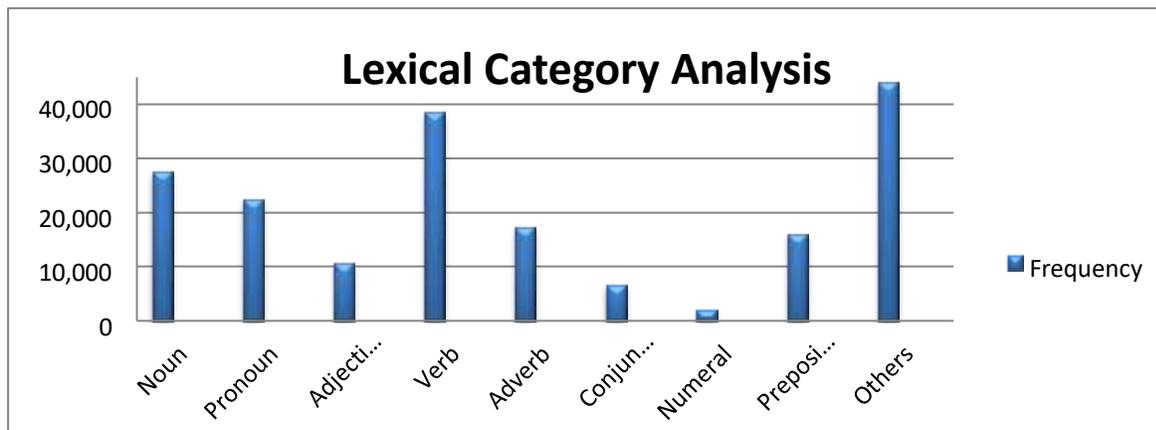


Chart 5.1 (a) showing the total number of frequencies of tokens in DNC

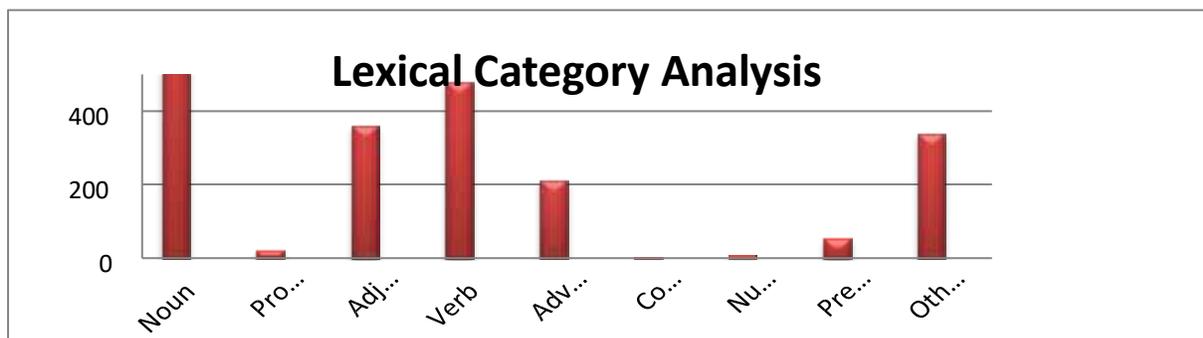


Chart 5.1 (b) shows the types of tokens in the DNC.

5.2 Adjectives as Sentiment Representative:

As mentioned in the study, words that qualify nouns are called adjectives. They can be of many types. In this paper, adjectives were used as representatives of sentiments. Findings related to adjectives are discussed below:

1. In the DNC, 360 types of adjective tokens were found, with a total frequency of 10,574 tokens.
2. In the DNC, the most frequent and the least frequent adjectives were found.
3. In the most frequent category, we have good, bad, much, other, happy, long, more, same, and few, etc.
4. In the least frequent category, we have bitter, quiet, rational, ready, gay, future, annoyed, proper, funny, and detached, etc.
5. Chart 6.2(a) shows the frequencies of sentiment token types.
6. Chart 6.2(b) shows some high-to-low-frequency adjectives found in DNC.

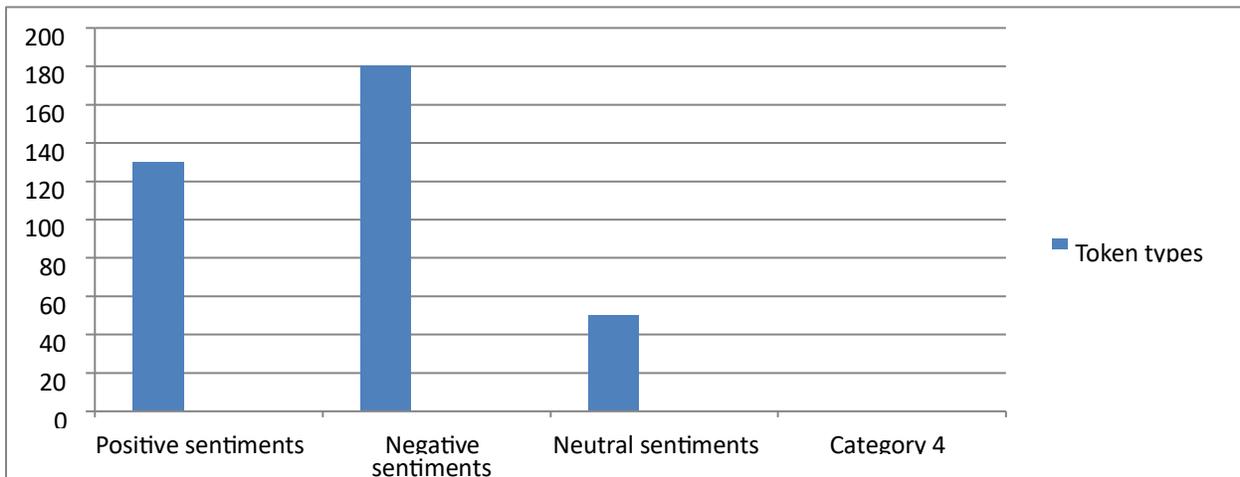


Chart 5.2(a) shows the number of sentiment token types.

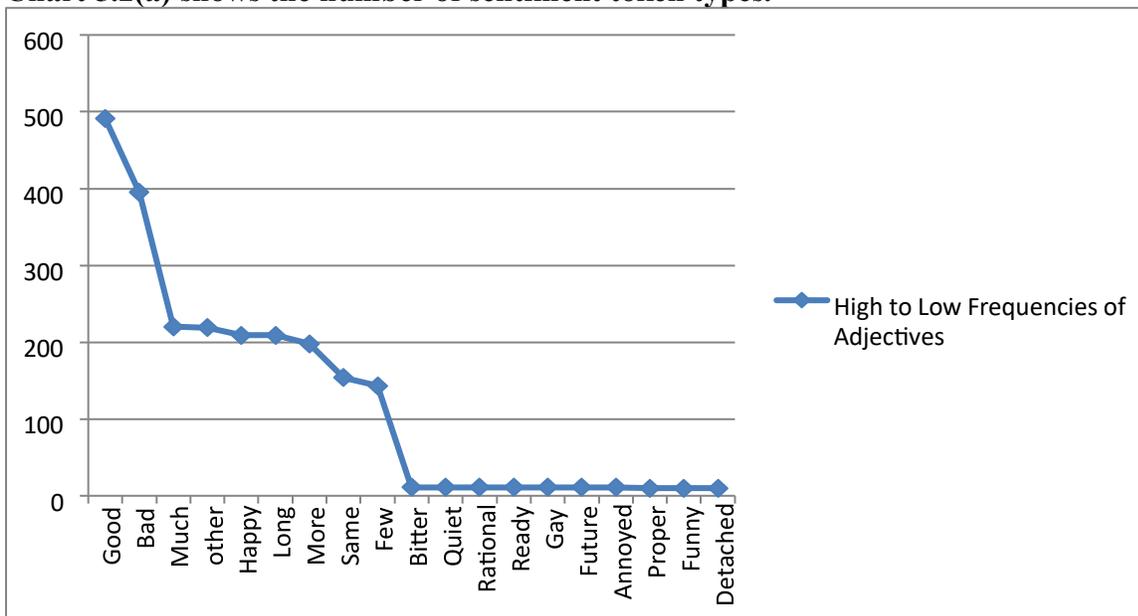


Chart 5.2(b): High to low frequencies of Adjectives

5.2.1 Sentiment Analysis:

Sentiment refers to the emotional tone or attitude expressed in a piece of text, speech, or communication. It can be categorized as positive, negative, or neutral, reflecting the author's

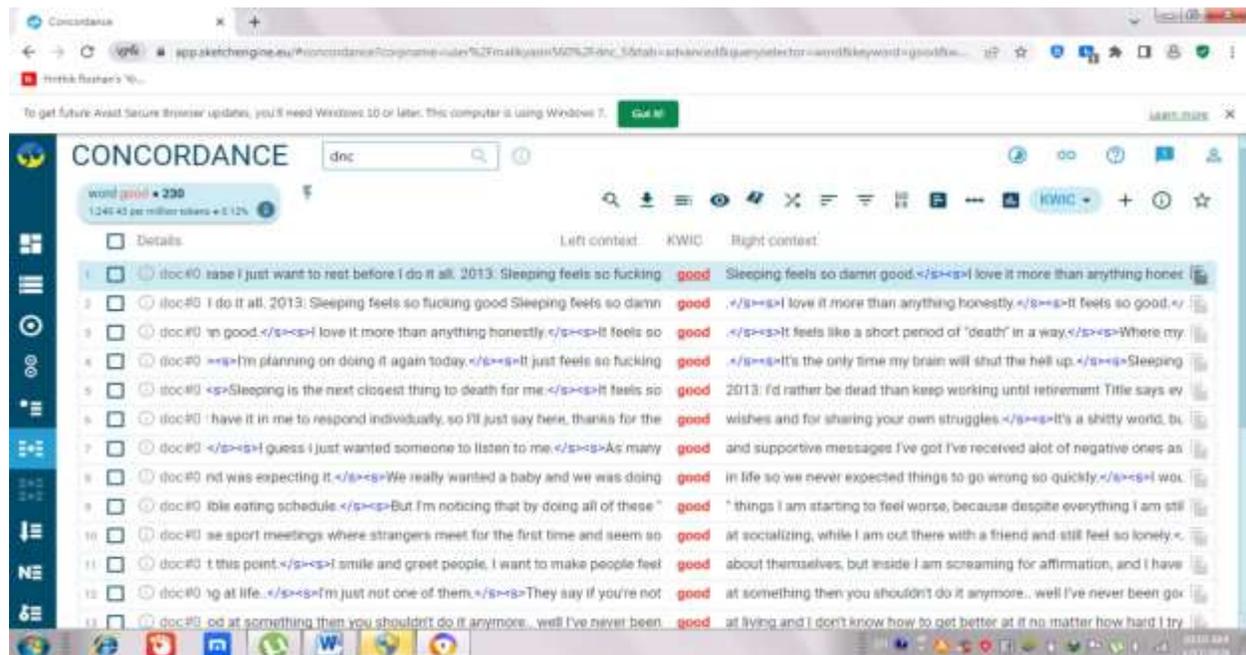
subjective feelings, opinions, or reactions toward a particular subject or topic. Sentiment analysis aims to determine and quantify sentiment in given content automatically. Sentiment analysis is a natural language processing technique used to determine the sentiment expressed in a piece of text, such as positive, negative, or neutral.

5.2.1.1 Positive Sentiments:

Positive sentiment is a favourable attitude or emotion expressed towards a topic or object. In this study, it was found that depression notes also consist of many adjectives, which can be considered as a positive sentiment analyser. In the DNC, 130 positive-sentiment representative adjective types were found. Some of them are good, beautiful, successful, warm, confident, valuable, pretty, cute, curious, intense, faithful, happy, great, emotional, supportive, special, like, regular, possible, mature, enjoyable, early, positive, easy, flexible, able, normal, self, basic, obvious, etc. Here are a few positive sentiments described below to show their usage.

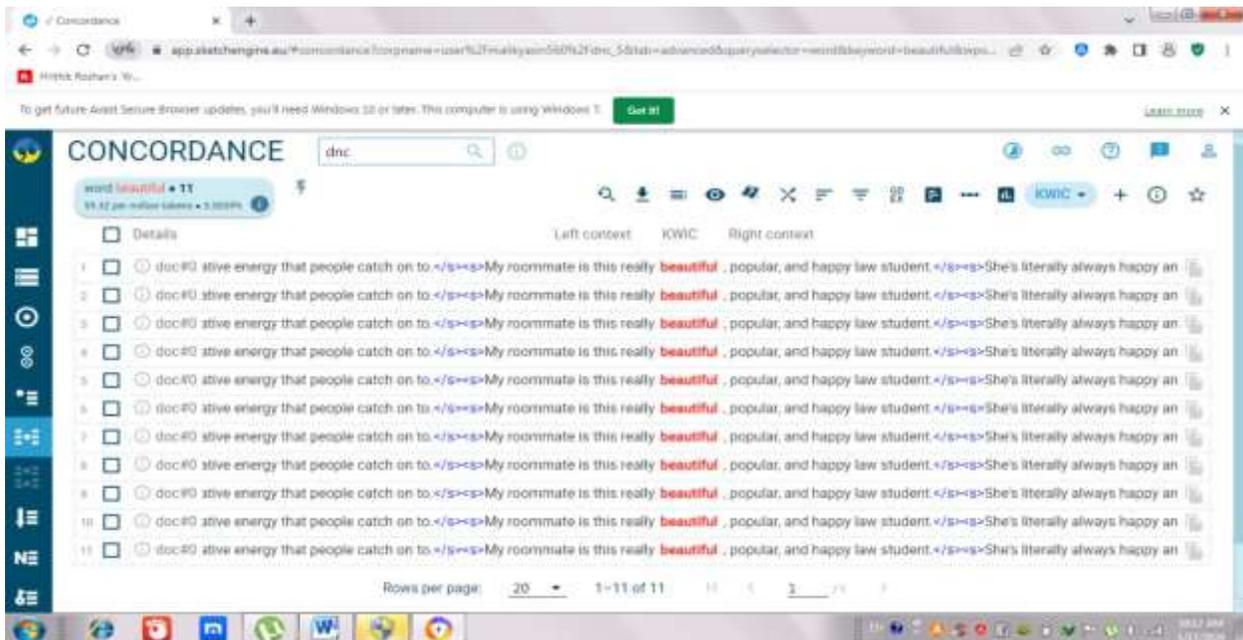
i. Good:

The word “good” is an adjective which shows positivity. Everyone who listens to or reads this word gets positivity or feels it. Even in the depression notes, it shows positivity. Depressed people mainly write that they feel good while sleeping. They want to get rid of depression to move towards goodness and mental peace.



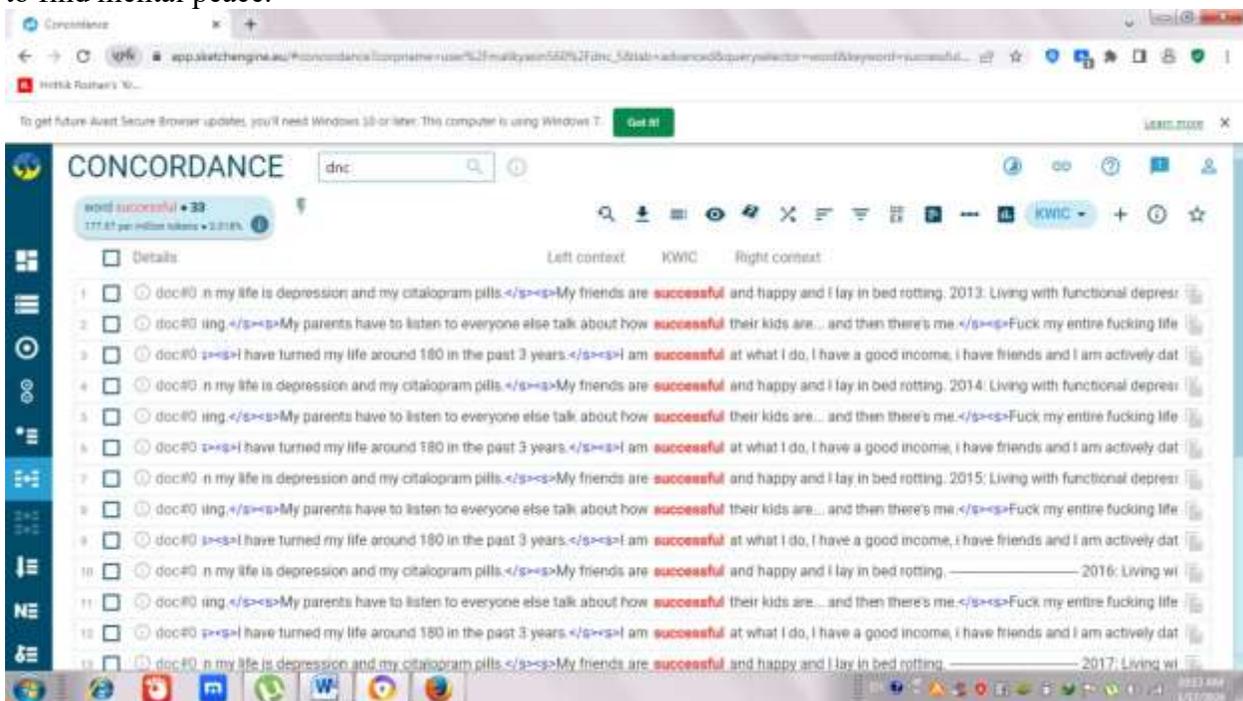
ii. Beautiful:

The word “beautiful” is an adjective which shows positivity. Everyone who listens to or reads this word gets positivity or feels it. Even in depression notes, Redditors use it to glorify something. Either they talk about their lacks, lost things like love or wealth, or about their beautiful memories. They want to get back every beautiful thing and memory.



iii. **Successful:**

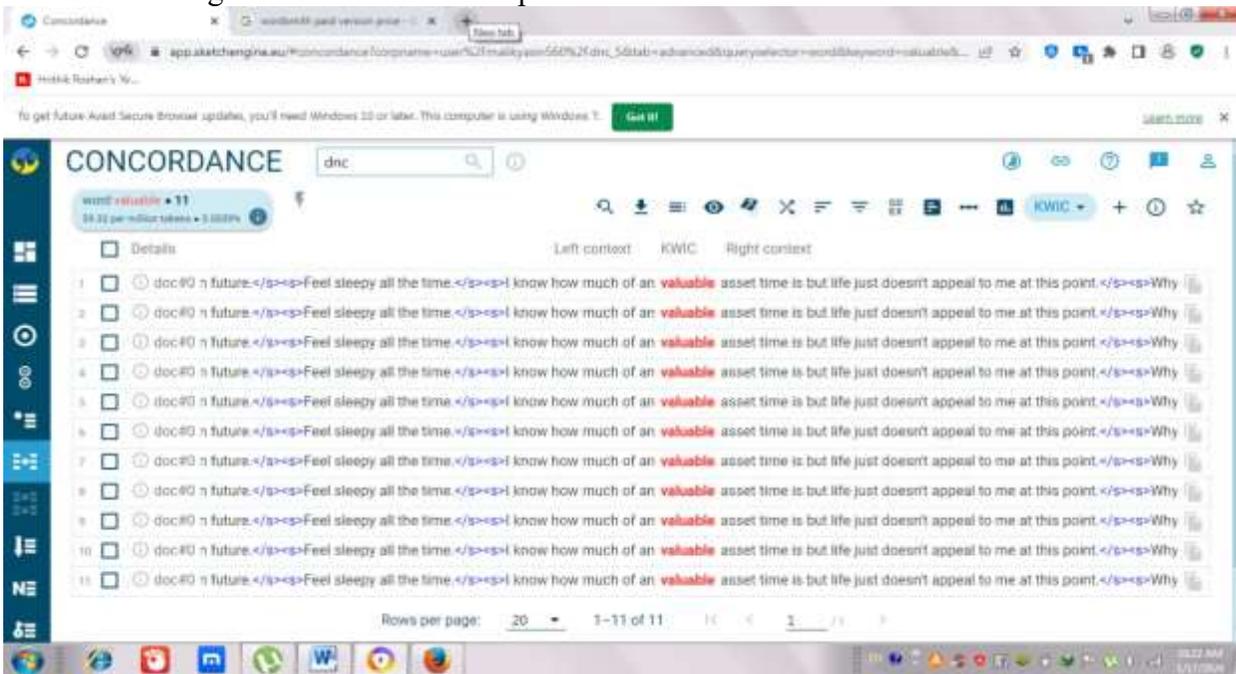
The word “successful” is an adjective which shows positivity. Everyone who listens to or reads this word gets positivity or feels it. Even in depression, this word conveyed some positivity, as Redditors wanted to share some positive memories with readers. They want to satisfy their souls to find mental peace.



iv. **Valuable:**

The word “valuable” is an adjective which shows positivity. Everyone who listens to or reads this word gets positivity or feels it. Even in depression notes, depressed people talk about their

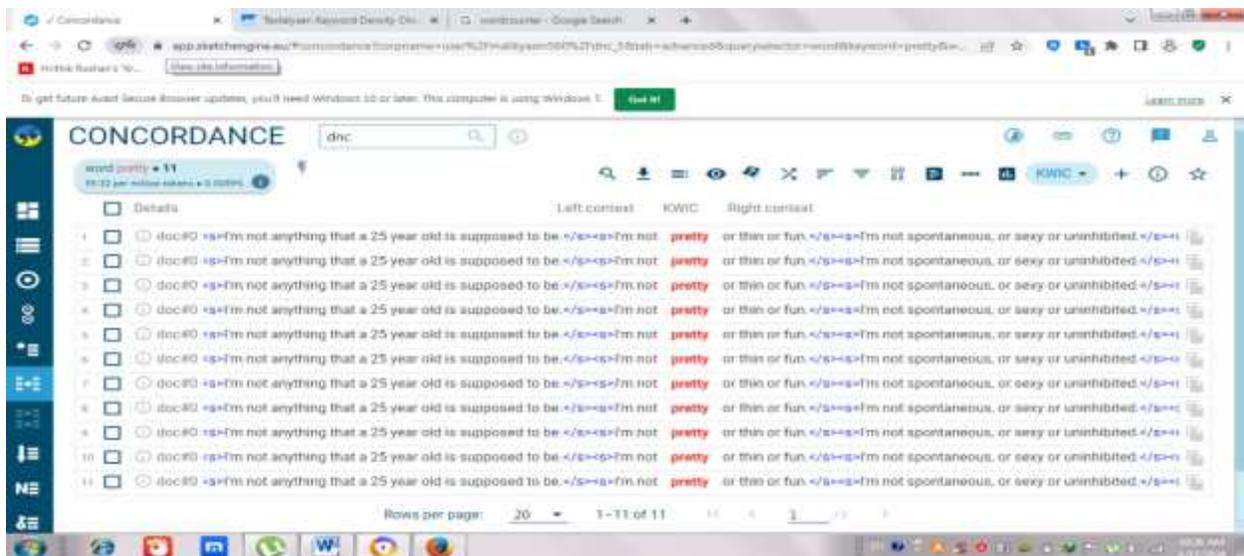
valuable things in life, and they become depressed about those things. They want to get their most valuable things back to attain mental peace.



v.

Pretty:

The word “pretty” is an adjective which shows positivity. Everyone who listens to or reads this word gets positivity or feels it. Even in depression notes, depressed people talk about the pretty things in life, and they become depressed about those things. They want to get their most prized possessions back to attain mental peace.

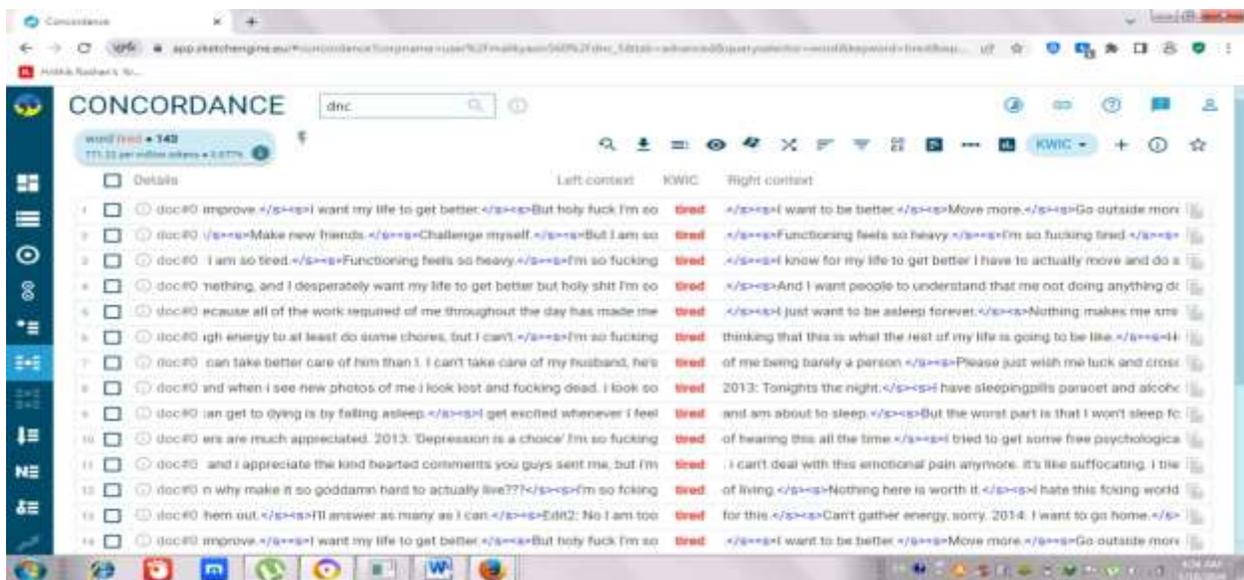


5.2.2 Negative Sentiments

Negative sentiment is an unfavourable attitude or emotion expressed towards a topic or object. It was found that depression notes also consist of many adjectives which can be considered as negative sentiment analyzer. In the DNC, 180 negative-sentiment representative adjective types were found. Some of them are tired, old, depressed, stupid, suicidal, sad, anxious, painful, hopeless, sick, sorry, impossible, guilty, awkward, relentless, meaningless, ill, weird, sightless, depressive, irrelevant, lazy, weak, losers, alcoholic, fake, paralyzed, stressful, poor, damaging, etc.

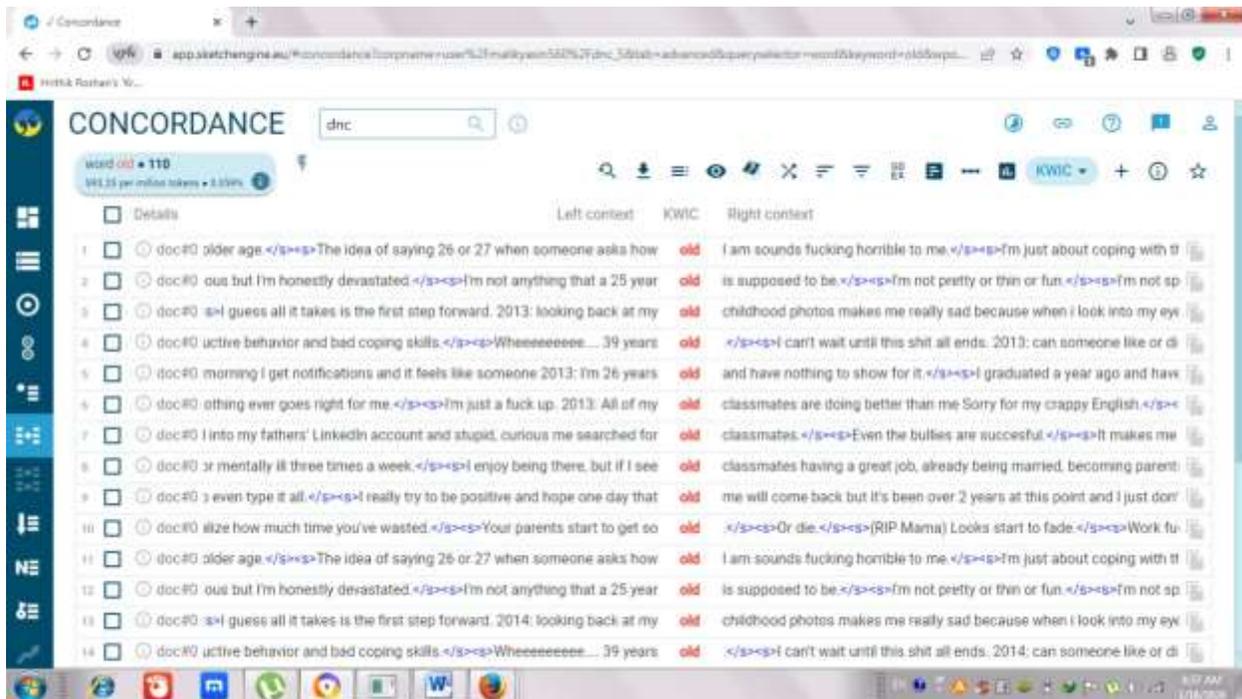
1. Tired:

The word “tired” is an adjective which shows negativity. Everyone who listens to or reads this word gets negative vibes or feels it. Even in depression notes, depressed people talk about things in life they are tired of, and they become depressed about those things. They want to get rid of things to attain mental peace.



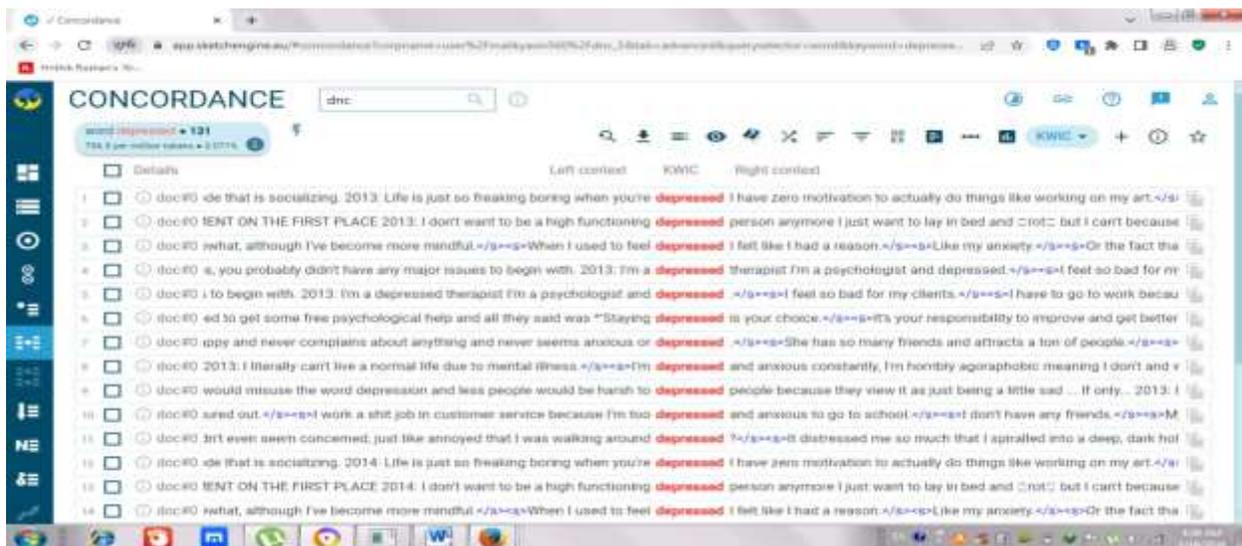
2. Old:

The word “old” is an adjective which shows negativity. Everyone who listens to or reads this word gets negative vibes or feels it. Even in depression, depressed people talk about things in life that are old in their lives, and they become depressed about those things. As they grow older day by day, the charm of life begins to diminish. After that, they also want to end their lives because of getting older and older. They want to get rid of things to attain mental peace.



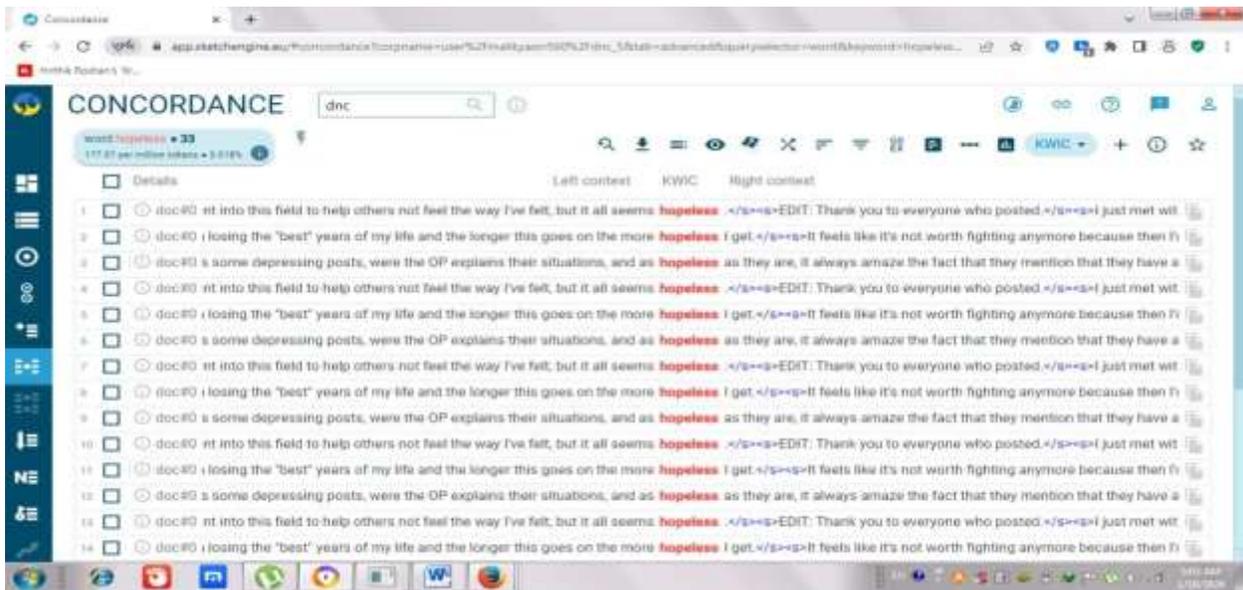
3. Depressed:

The word “depressed” is an adjective which shows negativity. Everyone who listens to or reads this word gets negative vibes or feels it. Even in depression notes, depressed people talk about things in life which are the cause of their depression, and they become depressed about those things. They want to get rid of things to attain mental peace.



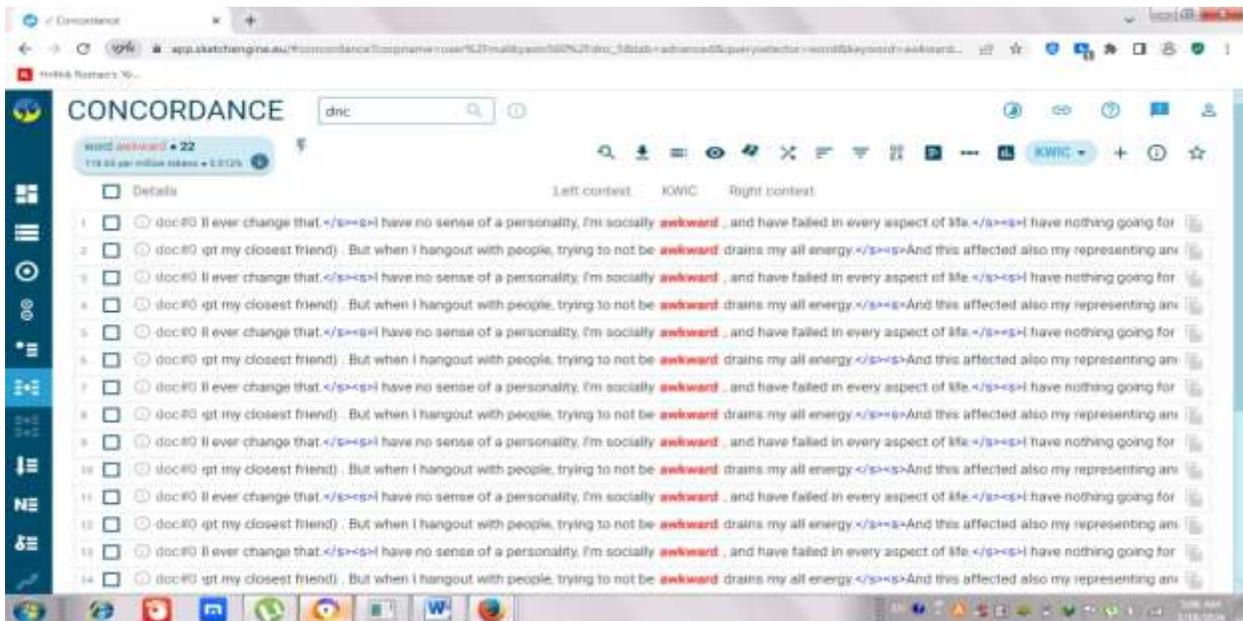
4. Hopeless:

The word “hopeless” is an adjective which shows negativity. Everyone who listens to or reads this word gets negative vibes or feels it. Even in depression, depressed people talk about their lives, which makes them feel hopeless, and they become depressed about those things. They want to get rid of things to attain mental peace. Sometimes, they want to get rid of life.



5. Awkward:

The word “tired” is an adjective which shows negativity. Everyone who listens to or reads this word gets negative vibes or feels it. Even in depression notes, depressed people talk about things in life they are tired of and become depressed about those things. They want to get rid of things to attain mental peace.

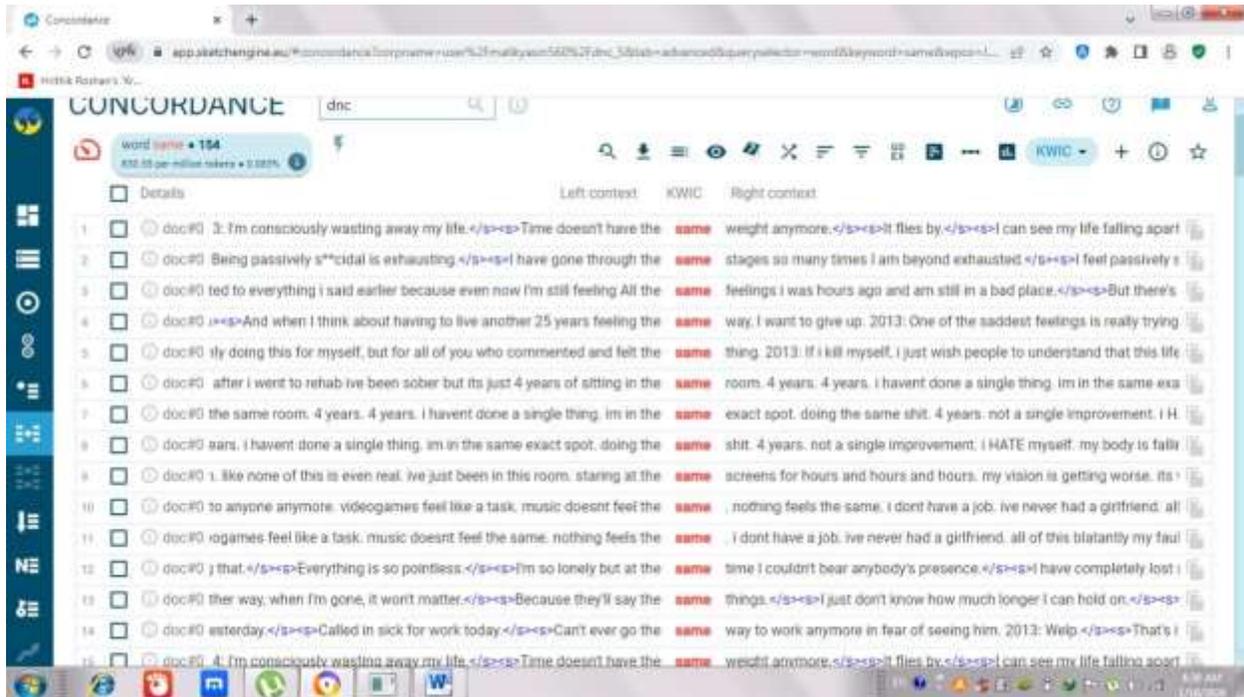


5.2.3 Neutral Sentiments:

A neutral sentiment is neither a favourable nor an unfavourable attitude or emotion expressed towards any topic or thing. In this study it was found that depression notes also consist of many adjectives which can be considered as neutral sentiment analyzer. In the DNC, 50 types of neutral sentiment representative adjectives. Some of them are others, a few, etc.

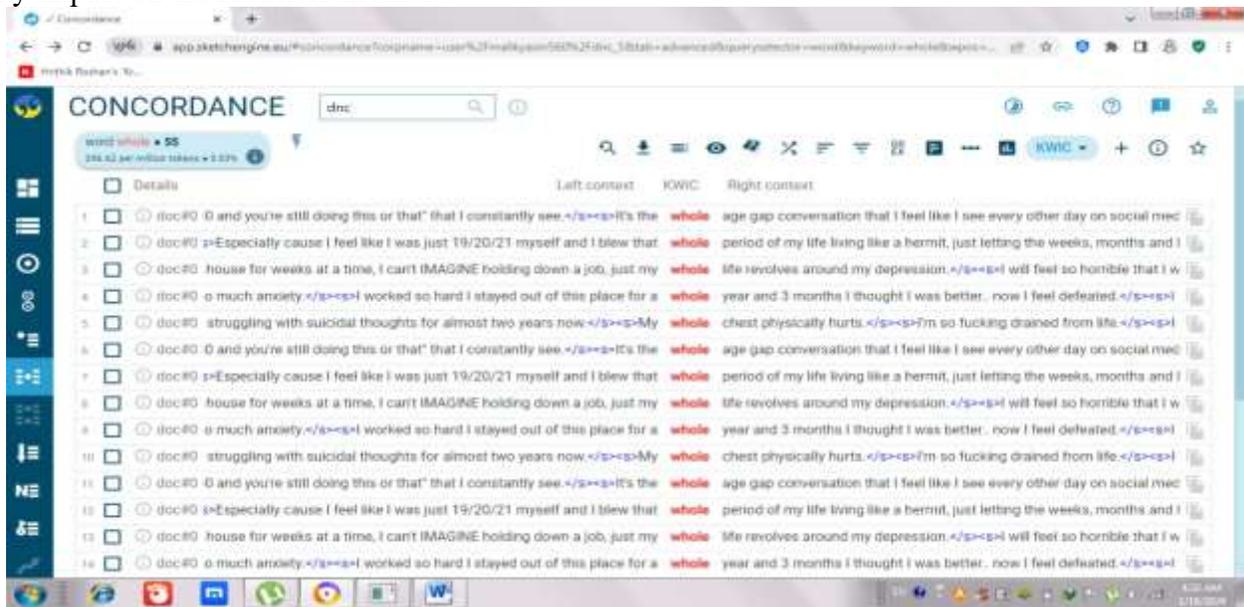
1. Same:

The word “same” is an adjective that shows neutral behaviour towards something. Everyone who listens to or reads this word gets neutral vibes or feels it. Its presence in the writing does not create any impact on the reader or listener.



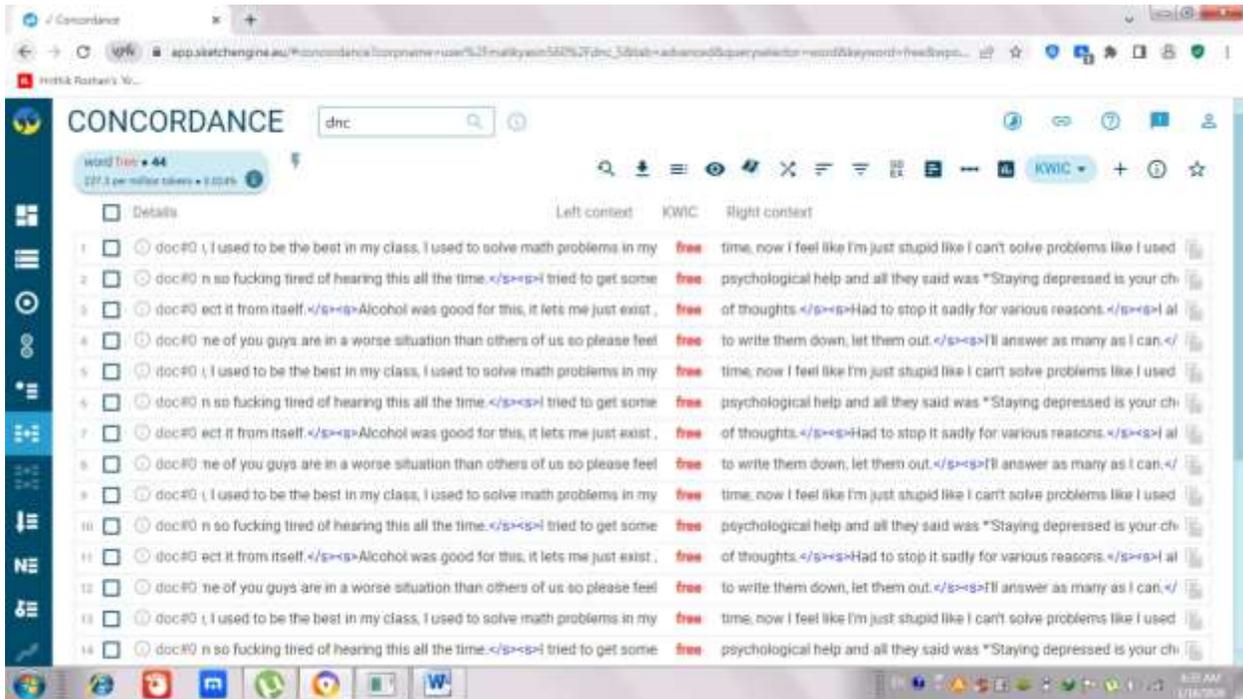
2. Whole:

The word “whole” is an adjective that shows neutral behaviour towards something. Everyone who listens to or reads this word gets neutral vibes or feels it. Its presence in the writing does not create any impact on the reader or listener.



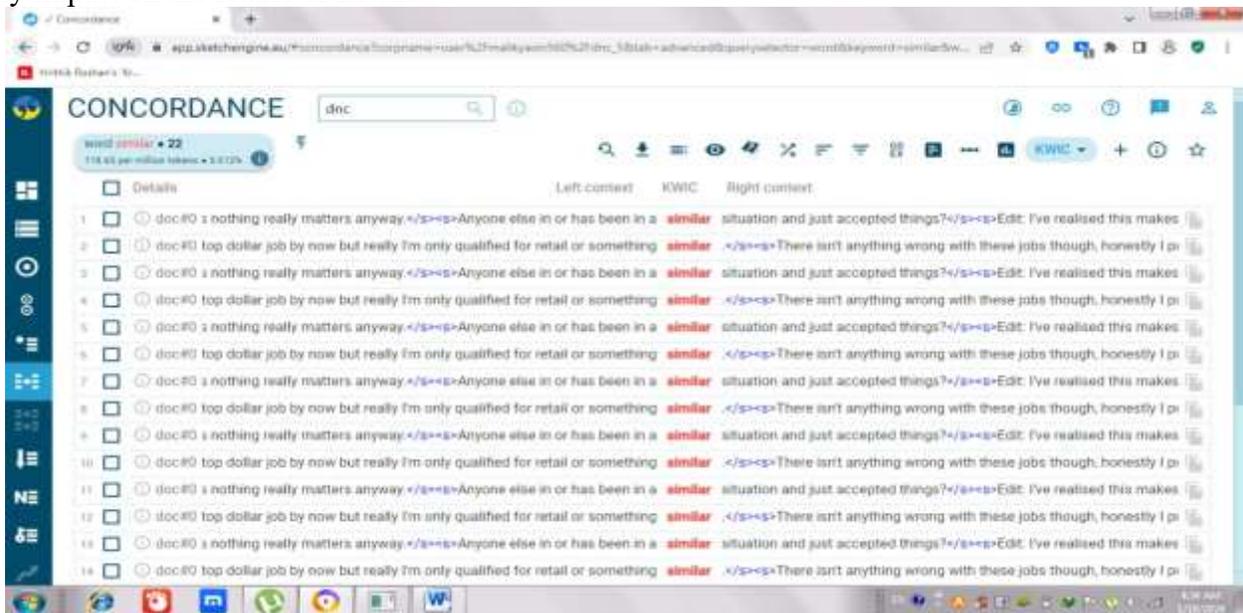
3. Free:

The word “free” is an adjective that shows neutral behaviour towards something. Everyone who listens to or reads this word gets neutral vibes or feels it. Its presence in the writing does not create any impact on the reader or listener.



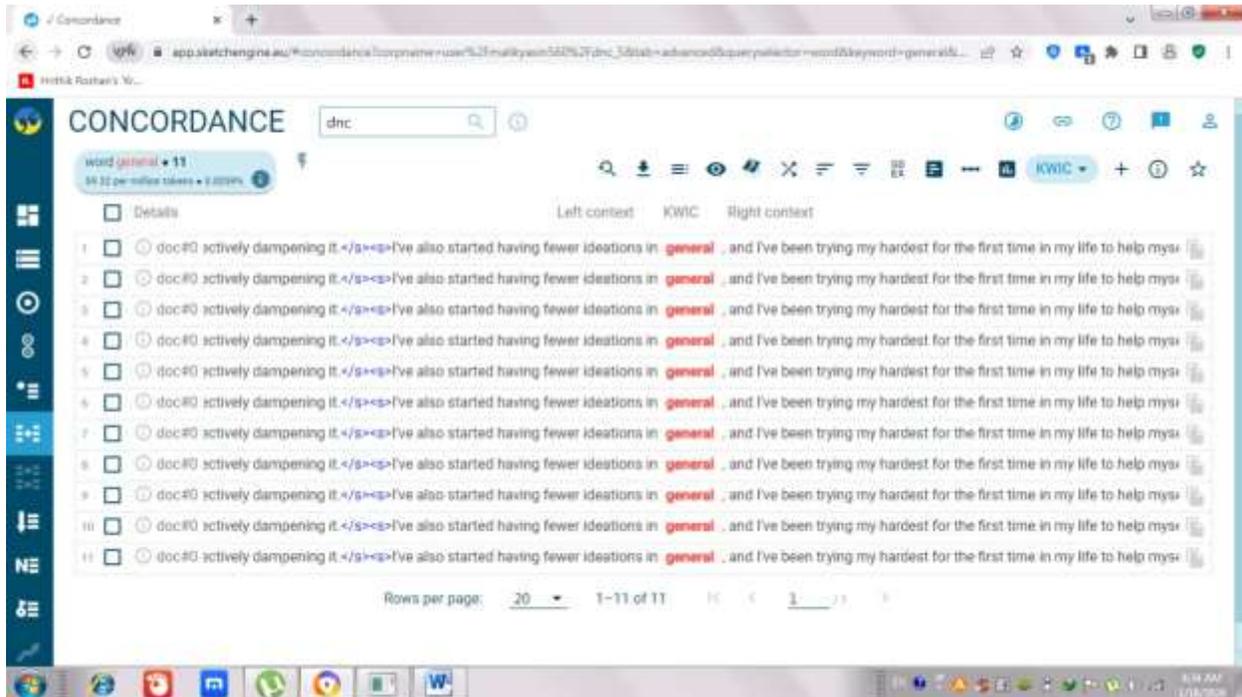
4. Similar:

The word “similar” is an adjective that shows neutral behaviour towards something. Everyone who listens to or reads this word gets neutral vibes or feels it. Its presence in the writing does not create any impact on the reader or listener.



5. General:

The word “general” is an adjective that shows the neutral behaviour of people towards something. Everyone who listens to or reads this word gets neutral vibes or feels it. Its presence in the writing does not create any impact on the reader or listener.



6. DISCUSSION:

In this article, according to the objectives of the study, all the lexical categories were found which were used by depressed Redditors in the years 2013 to 2023. The lexical categories were examined which were used in depression notes in order to analyze the choices made by depressed people. In this study adjectives were considered as sentiment analyzers. In this study, the discussion of lexical categories and their types indicates that a total of 2280 token types were found in the DNC. A total of 794/2280 noun token types were found in DNC, the highest number. Then, a total of 480/2280 verb token types were found in DNC. Then, a total of 360/2280 adjective token types were found in DNC. Then we found 339/2280 token types for other entities, including punctuation and contractions. Then, 212/2280 adverb token types were found. Then, 55/2280 token types of prepositions. Then we found 23/2280 token types of pronouns. Then we found 10/2280 token types of numerals. Moreover, we found 360/2280 token types of conjunctions. Despite this, if we consider the frequencies of lexical categories, there were 1,85,420 total tokens across lexical categories in DNC. The most frequent lexical category was verbs, with a total of 38,576 tokens in DNC. After verbs, the most frequent category was nouns, with 27,679 tokens in DNC. After nouns, the most frequent category was pronouns, with a total of 22,530 tokens in DNC. After pronouns, the most frequent category was adverbs, with a total of 17,283 tokens in DNC. After adverbs, the most frequent category was prepositions, with 16,042 tokens in DNC. After prepositions, the most frequent category was adjectives, with a total of 10,574 tokens in DNC. The least frequent category was conjunction, with a total of 6,682 tokens

in DNC. Except for these, we found a total of 2050 numerals in DNC. We also found other tokens that consist of punctuation, contractions, and interjections, with a frequency of 44,044 tokens in DNC. Depressed people use more and more verbs to show that they want to perform some sudden action in order to get rid of depression. They use adjectives to convey their emotions and sentiments. They use adverbs to show how they want to perform the actions. They use nouns to glorify the things or materials which they need. They use the pronoun I the most because they always talk about themselves. They use prepositions to show the positions of materials and things. They use numerals to show an address or a time. They use different conjunctions to connect their posts, making these writings more meaningful. They use contractions to reflect their dialect and language use. Despite everything, the word choices of depressed people create an impact on the reader's mind, and sympathy bears from their hearts.

In this study, adjectives have been used as sentiment-analysis entities. The researchers identified 360 adjective token types in DNC, comprising 180 negative, 130 positive, and 50 neutral tokens. As they made the study and analyzed the depression notes, they found that depressed people are more sensitive than the tension-free people. They have many reasons for being depressed. These reasons could include financial losses, losing games, people's disloyalty, being left by others, societal disrespect, and facing various social barriers. Depressed people are so hopeless that they possibly want to end their lives. They just need a loyal friend or someone to release their depression. They keep themselves in a state of depression, which is not under their control. They keep talking about the reasons why they all bear these circumstances. In the depression notes, the researchers found all three, positive, negative and neutral sentiments showed by the selection of different choices of adjectives. Positive sentiments include good, beautiful, successful, warm, confident, valuable, pretty, cute, curious, intense, faithful, happy, great, emotional, supportive, special, like, regular, possible, mature, enjoyable, early, positive, easy, flexible, able, normal, self, basic, obvious, etc. These all show positivity. Negative sentiments include tired, old, depressed, stupid, suicidal, sad, anxious, painful, hopeless, sick, sorry, impossible, guilty, awkward, relentless, meaningless, ill, weird, sightless, depressive, irrelevant, lazy, weak, loser, alcoholic, fake, paralyzed, stressful, poor, damaging, etc. All these show negativity. Neutral sentiments include other, few, constant, etc. All these show neutral behaviour. All positive sentiments create a positive impact on others; negative sentiments create a negative impact, and neutral sentiments create a neutral impact on other people's minds.

6. IMPLICATIONS AND CONCLUSION:

Implications of this study can be: people could be able to read about the depressed community and provide them with emotional support. Another implication is that there will be development opportunities. This study concludes that depressed people use all the categories of words, but they mostly use verbs in order to show their purposive action, which they want to show. They use positive, negative and neutral adjectives to show their sentiments. Their made choices have all positive, negative, and neutral impacts on others' sentiments. A transformation can be made in their lives by showing them sympathy through responding to their posts and notes. The transformation is not just a task of researchers; it is the task of every individual in society. As it is widely known that society moulds a man, the depressed community can be moulded into a happier one with the support of the people of society. Everyone should be allowed to express their sentiments to others without being judged or mocked. To build a healthier society, a cooperative

community plays a crucial role. Sentiment analysis shows that if emotional support is given to depressed people, there can be a huge transformation in society.

REFERENCES

1. Adda, G., Mariani, J., Lecomte, J., Paroubek, P., & Rajman, M. (1998). The GRACE French part-of-speech tagging evaluation task. In A. Rubio, N. Gallardo, R. Castro, & A. Tejada (Eds.), *Proceedings of the First International Conference on Language Resources and Evaluation (LREC)* (Vol. 1, pp. 433–441).
2. Alpaydin, E. (2004). *Introduction to machine learning (Adaptive computation and machine learning series)*. MIT Press.
3. Anthony, J. H. (2007). *Probability and statistics for engineers and scientists* (3rd ed.). Duxbury Press.
4. Dave, K., Lawrence, S., & Pennock, D. M. (2003). Mining the peanut gallery: Opinion extraction and semantic classification of product reviews. In *Proceedings of the 12th International World Wide Web Conference (WWW, 2003)* (pp. 519–528). ACM. <https://doi.org/10.1145/775152.775226>
5. Jansen, B. J., Zhang, M., Sobel, K., & Chowdury, A. (2009). Micro-blogging as online word of mouth branding. In *Proceedings of the 27th International Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '09)* (pp. 3859–3864). ACM. <https://doi.org/10.1145/1520340.1520599>
6. Lafferty, J. D., McCallum, A., & Pereira, F. C. N. (2001). Conditional random fields: Probabilistic models for segmenting and labelling sequence data. In *Proceedings of the 18th International Conference on Machine Learning (ICML 2001)* (pp. 282–289). Morgan Kaufmann.
7. Manning, C. D., & Schütze, H. (1999). *Foundations of statistical natural language processing*. MIT Press.
8. Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1–2), 1–135. <https://doi.org/10.1561/1500000011>
9. Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up? Sentiment classification using machine learning techniques. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2002)* (pp. 79–86). Association for Computational Linguistics.
10. Pedersen, T. (2000). A simple approach to building ensembles of naive Bayesian classifiers for word sense disambiguation. In *Proceedings of the 1st North American Chapter of the Association for Computational Linguistics Conference (NAACL 2000)* (pp. 63–69). Morgan Kaufmann.
11. Read, J. (2005). Using emoticons to reduce dependency in machine learning techniques for sentiment classification. In *Proceedings of the ACL Student Research Workshop* (pp. 43–48). Association for Computational Linguistics.
12. Schmid, H. (1994). Probabilistic part-of-speech tagging using decision trees in *Proceedings of the International Conference on New Methods in Language Processing* (pp. 44–49).
13. Shannon, C. E., & Weaver, W. (1963). *The mathematical theory of communication*. University of Illinois Press.
14. Wilson, T., Wiebe, J., & Hoffmann, P. (2005). Recognizing contextual polarity in phrase-level sentiment analysis. In *Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing (HLT/EMNLP 2005)* (pp. 347–354). Association for Computational Linguistics.
15. Yang, C., Lin, K. H.-Y., & Chen, H.-H. (2007). Emotion classification using web blog corpora. In *Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence (WI, 2007)* (pp. 275–278). IEEE Computer Society.