Recommender System for Postpartum Depression Monitoring based on Sentiment Analysis

Marcílio B. Carneiro[1], Mário W. L. Moreira[1], Silas S. L. Pereira[1], Erica L. Gallindo[1],
and Joel J. P. C. Rodrigues[2,3]
[1]*Federal Institute of Education, Science, and Technology of Ceará (IFCE), Aracati, CE, Brazil*
[2]*Federal University of Piauí (UFPI), Teresina, PI, Brazil*
[3]*Instituto de Telecomunicações, Portugal*
marcilio.bc@gmail.com, {mario.wedney, silas.santiago, erica.gallindo}@ifce.edu.br
joeljr@ieee.org

2020 IEEE International Conference on E-health Networking, Application & Services (HEALTHCOM) | 978-1-7281-6267-6/21/$31.00 ©2021 IEEE | DOI: 10.1109/HEALTHCOM49281.2021.9398922

***Abstract*—Emotions influence all aspects of human behavior. All of these aspects shape people's lives, directly impacting their ways of life. Some diseases are directly linked to emotions. Among them, depression is one of the diseases with the greatest impact on society. Hence, faced with this problem, the objective of this study is to present a context-aware solution based on text mining for gestational depression prevention. This system uses text mining to analyze documents filled from pregnant women in order to identify their feelings through natural language processing techniques and probabilistic algorithms. As a case study, the analyzed texts were obtained from forms answered by pregnant women. The model performance is evaluated using metrics associated with the confusion matrix. The results show that the proposed model has achieved a reliable performance in all metrics, mainly when classifying new cases. Thus, the results obtained by the model can be used as support to health professionals in monitoring high-risk pregnancies.**

***Index Terms*—recommender systems, text mining, natural language processing, sentiment analysis, pregnancy**

I. Introduction

Emotions modify human behavior. Feelings such as anger, happiness, sadness, surprise, disappointment, among others, are some emotions that people have daily. These emotions directly contribute to human behavior. Factors such as unemployment, violence, financial situation, goals at work, overstudying, among others, are situations capable of generating an emotional imbalance, causing several diseases, and depression is the main one. According to the World Health Organization (WHO), depression affects 300 million people worldwide. Thus, making this disease more disabling by 2020. According to [1], the entire emotional process plays a decisive role in human conduct. Depression is a disease that has a considerable impact on the patient's life and his family. This is characterized by a change in mood, a high degree of sadness, and a lack of interest in daily activities.

Gestational depression is a disorder that can harm the mother-child relationship. Hence, the importance of the mother's physical and psychological well-being throughout her pregnancy and after it. Especially in the period immediately before the childbirth and in the postpartum. It is common for the mother to go through a prolonged period of sadness in the postpartum due to hormonal changes, named baby blues and maternity blues. Numerous symptoms can indicate depression. According to [2], these symptoms include irritability, frequent crying, feelings of helplessness and hopelessness, lack of energy and motivation, sexual disinterest, eating and sleeping disorders, and the feeling of to be incapable to deal with new situations, as well as complaints psychosomatic. These situations occur due to countless changes in a short period of a woman's life. Several factors influence gestational depression, such as the lack of family support and financial difficulties. Thus, the health professional must be attentive to the patient, who needs, in these cases, daily monitoring.

Besides, the emotions shared among people can also be found within texts, images, and sounds. Sentiment analysis is the study of these emotions. Due to the growth of social networks and the countless data generated, sentiment analysis has been of great use to identify people's emotions and reactions. Greco and Polli define text mining as a knowledge-intensive process in which a user interacts with a large number of documents using analysis tools [3]. The goal is to extract useful information from collections of documents. This information is identified by patterns of interest in unstructured textual data.

Smart systems and devices are capable of providing important data on human emotions and have been widely used in the health area. For this purpose, data can be used to monitor patients at risk of developing mental illness. In affective computing, it is studied the way that computers can recognize, model, and respond to human emotions (among other aspects) and, in this way, the form they can express them through a computational interface/interaction using sentiment analysis to identify emotions [4].

Based on the discussion so far, this paper presents the design and implementation of a smart emotion-aware solution based on text mining for gestational depression prevention, with the objective of supporting health professionals to identify the patient's emotions, early diagnosing and avoiding gestational depression. The main purpose is to identify changes in the patient's mood during the gestational period, allowing the professional to acquire more information about the patients to facilitate their diagnosis. The solution is based on the theories of Paul Ekman, an American psychologist and pioneer in the study of emotions, which relating some feelings that can be

978-1-7281-6267-6/21/$31.00 ©2021 IEEE

identified through text mining.

The remainder of the paper is organized as follows. Section II presents the related works, comparing their several approaches. In Section III, the proposed prototype is presented, characterizing the entire process necessary for the development of the predictive model. Section IV presents the results obtained from the experiments described in the previous section. Finally, Section V presents the final considerations of the study and suggestions for future work.

II. Related Work

A. Using a Text Mining Tool to Support Text Summarization

In [5], a graph representation method is proposed to help students in tasks like writing abstracts and essays using a text-mining tool. The Sobek application is based on a development based on a distance algorithm. This algorithm extracts relevant terms from the analyzed text. The authors divided the writing process into three phases. In the first, called pre-writing, the students reads the text that will be summarized by them and identifies the ideas necessary for writing. After this first reading, the student uses the Sobek to extract relevant terms and relationships from the text. Analyzing the graphical representation of the text, the student organizes his ideas and reads the text once more, this time based on the extracted terms and their relations. In the next phase, the writing phase, the student uses the text graphic as a comparison, in relation to the abstract, to ensure that it is in accordance with the main ideas of the text. In the third and final phase, called rewriting, the student reviews the text already structured and written, facilitating the completion of his work. The effectiveness of the experiment was verified through a study with twenty high school students between 15 and 18 years old in a computer lab. According to the results of the experiment, the students used 61.6% of the terms extracted from the Sobek, showing its relevance in helping the student in the construction of abstracts and essays through text mining.

B. Mining the World's Interest through Emotion Analysis

Saravia *et al.* proposed a tool for understanding the interests and emotions of users in a worldwide proportion [6]. The Emoviz system is a visualization platform that consists of identifying a possible interest or trend through emotions. The Emoviz users can insert a certain word and check for the global interest in the search term. The architecture is composed of four phases. In the first phase, the preprocessing phase, non-relevant words are eliminated. Only words of interest from Twitter are extracted. In the second phase, techniques such as ruled-based extraction, keyword extraction, and part-of-speech tagging are presented, to contribute to creating the Emoviz word cloud. In the third phase, in which emotions are analyzed, the collection of tweets to their respective classes is labeled using an internal classification algorithm. In the last phase, the classification of interest, in which the output with the respective class of a particular word or phrase returns. The detailed architecture data set output is used to implement the visual components of EmoViz, such as maps, graphics, and a cloud of words. To address the problem, the authors used around 6,000 tweets to identify users' emotions regarding the iPhone 6 opening ceremony in 2014.

C. Chromotherapy and Affective Computing: Identifying Anxiety States

Muniz *et al.* proposed a Web tool, named CroCA, to simulate an online diary where users write their routines [7]. These texts are processed and, if there are traces of anxiety in the text, the CroCA tool interface changes its color, with a visual response in blue to calm and red to provoke excitement. The authors developed a chatbot capable of returning appropriate responses when interacting with the user. To develop the tool, the authors used affective computing procedures. Regarding colors to identify user behavior, the authors based themselves on scientific works that emphasize the effect of colors on the organism.

D. A Mobile Application to Support Puerperal Women with Postpartum Depression based on the Edinburgh Scale

Brito developed a mobile application to support the screening of postpartum women based on the Edinburgh Scale, which consists of a self-assessment instrument composed of 10 items related to depressive symptoms frequently observed in the puerperium [8]. The methodology adopted for the development of this research is divided into four stages, to know, bibliographic research and initial proposal, case study, application development, and testing. Results showed that with the use of the mobile application is possible to identify traces of postpartum depression in pregnant women accurately. For the case study, the author applied a questionnaire to psychologists to find out their contact with the puerperal woman and their acceptance of using mobile applications. When analyzing the data, it is clear that most psychologists attend puerperal women, thus being a relevant study for the interviewees. It is important to notice that 80% of the interviewees agreed to the application development. Another questionnaire was applied for a performance assessment of the Puerperium APP, in which eight mothers and two psychologists evaluated the application positively.

E. An Ensemble Neural Network Model for Benefiting Pregnancy Health Stats from Mining Social Media

In [9], the authors present an ensemble approach based on long short-term memory (LSTM) and CNNs combined with a support vector machine (SVM) classifier to detect pregnancy cases from Twitter data. The authors also present an exploratory sentiment analysis study on pregnant women accounts as one of the contributions of their research. From the data collected from Tweets, this study performed the identification and ranking of terms most commonly related to physical and behavioral conditions. The authors organized the pregnancy-associated sentiment in neutral, happy, distress, and annoyed. The authors applied sentiment-lexical feature co-occurrence using the log-likelihood ratio (LLR) method to identify trigger words for these sentiments. According to this research, the results indicated a direction to support the identification of pregnancy sentiment classes.

F. Sentiment Analysis Based on Deep Learning and Its Application in Screening for Perinatal Depression

In [10], the authors apply sentiment analysis in the context of perinatal depression screening scenarios, which is a time-consuming manual screening process since involves the examination of medical history, mental examination, psychological assessment, among other information. They built an LSTM model to extract and classify the user's emotional features. Besides, emotions were also considered an important feature. For the experiments, they used a dataset from survey data of the Edinburgh Postnatal Depression Scale questionnaire combined with the circle of user friends in the WeChat social network. The dataset was split with a holdout strategy, considering 80% for the training step. The results confirm the application of this deep learning neural network, as it was possible to achieve similar results for the Edinburgh Postnatal Depression Scale Screening.

Table I presents a comparison of the related works. For this comparison, this study took into account the following characteristics, to know, the use of emotion analysis techniques, the use of real data to generate the predictive model and probability computation as output.

TABLE I
COMPARISON OF THE CURRENT STATE OF THE ART APPLICATIONS.

Research	Emotion analysis	Real data	Probability
Reategui *et al.* [5]	√	√	
Saravia *et al.* [6]	√	√	
Muniz *et al.* [7]	√	√	
Brito [8]		√	
Seixas *et al.* [11]			√
Carneiro *et al.*	√	√	√

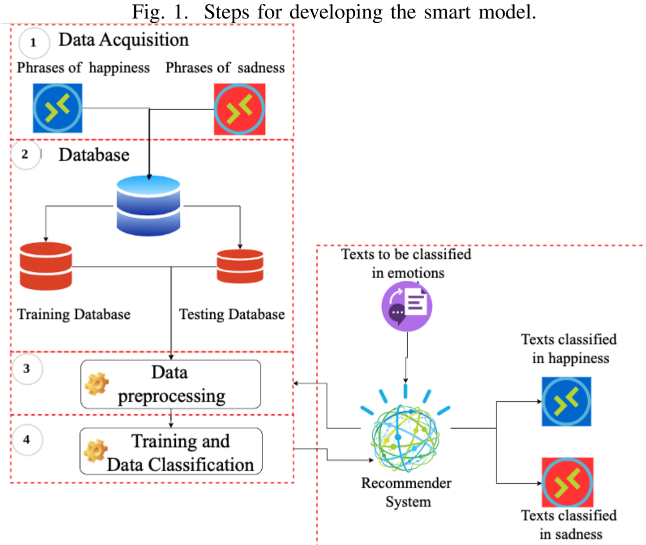
III. A RECOMMENDER SYSTEM FOR POSTPARTUM DEPRESSION MONITORING

Figure 1 shows the execution flow steps to obtain the proposed predictive model. In the first step, the sentences are collected from forms filled by pregnant women. Then, in the second step, the database is organized, being divided into a training and testing base. In the third stage, natural language processing (NLP) techniques are used for preprocessing the text data bases. In the next stage, data training is performed to obtain a supervised classification model. In the fifth stage, it is shown whereby the data are classified and, finally, the last stage presents the model performance in the classification of these texts concerning the emotions of happiness and sadness.

A. Data acquisition

In the data collection, two emotions were chosen among the six primary emotions studied by Paul Ekman. The emotions used were happiness and sadness. These emotions are quite distinct and a high success rate is expected from the model

Fig. 1. Steps for developing the smart model.

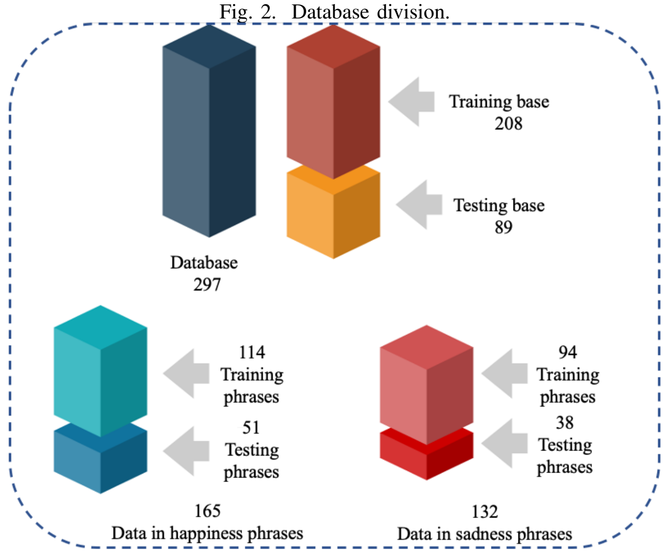


about its predictive classification. For the construction of the phrases of these emotions, pregnant women, mothers, nursing technicians, and a student of psychology at the Potiguar University in Brazil were recruited. Thus, eleven people were interviewed to construct the database.

B. Database description

The database consists of a text file containing 297 phrases, which 165 phrases belong to the happiness class and 132 phrases belong to the sadness class. The database was divided into two other bases, namely, the training base and the test base, as can be seen in Figure 2.

Fig. 2. Database division.



Considering the literature, the most common division for a database is the proportion of 70% for training and 30%

for testing. Table II describes the number of phrases used for training and testing.

TABLE II
DATA DIVISION IN TRAINING AND TESTING.

Base	Training	Testing
Complete base	208	89
Happiness base	114	51
Sadness base	94	38

The decision to choose this proportion is related to the number of records in the database, a number considered low. If this number were higher, a proportion of 80% and 20% or even 90% and 10% could have been used for training and testing, respectively. However, in this specific case, a very low number of testing data could compromise the model performance.

C. Data preprocessing

In this step, the data is processed using ML techniques. At this point, planning and processing are required to obtain favorable performance in the final process.

A tool from the Natural Language Toolkit (NLTK) library was used to perform the necessary techniques in the data preprocessing. According to [12], the NLTK platform is used to develop Python language programs that perform with human language data for applications based on NLP. According to the authors, the NLTK defines an infrastructure that can be used to build NLP programs in Python; provides basic classes to represent data relevant to the NLP; standard interfaces to perform tasks such as tokenization, part-of-speech, parsing and text classification; and standard implementations for each task that can be combined to solve complex problems.

The stopwords are words that are not relevant for understanding the emotions in the text. Some examples among others are “as”, “from”, “to”, “with”, “between”, and “was”. According to [13], the elimination of stopwords is a method applied exclusively to texts. Terms (words) that are prepositions, articles, conjunctions, can be considered irrelevant to the task regardless of the task to be performed. However, terms that are stopwords are language-dependent. Stopwords are very common and occur with high frequency in the language considered. Thus, the fact of eliminating them from the attribute vector causes a drastic reduction in the set of attributes. Table III presents what a phrase would look like when applying the stopwords technique.

TABLE III
APPLICATION OF STOPWORDS.

	Phrase
With stopwords	We are happy we have a lot of love.
Without stopwords	happy love.

To obtain the words, as seen in the table above, the command is used:

```
stopwordsnltk = nltk.corpus.stopwords.words(english)
```

Then, the stopwords are removed by applying the function below.

```
def removestopwords(text):
    phrases = []
    for(words, emotion) in text:
            semstop = [p for p in words.split()
            if p not in stopwordsnltk]
            phrases.append((semstop, emotion))
    return phrases
```

Another NLP technique used was stemming. This technique allows the elimination of prefixes and suffixes, reducing the representative spaces of the data considerably. Table IV presents the stemming technique application.

TABLE IV
STEMMING APPLICATION.

	Phrase
Without stemming	happy, loving, happiness, sad.
With textit stemming	happ, lov, happ, sad.

The stemming technique can also be found on the NLTK platform and can be applied as follows.

```
def applystemmer(text):
    stemmer = nltk.stem.RSLPStemmer()
    phrasesestemming = []
    for (words, emotion) in text:
        cste = [str(stemmer.stem(p))
        for p in words.split() if p not in
        stopwordsnltk]
        phrasesestemming.append((cste, emotion))
return phrasessestemming
```

D. Training and data classification

An important tool in this study is the naive Bayes algorithm, which performed the training and classifying tasks of data. According to Hasan *et al.*, the naive Bayes classifier is a simple and frequently used method for supervised learning, providing a flexible way to deal with any number of attributes or classes. Besides, small amounts of bad data do not hinder the performance results [14].

Consider a random variable C that denotes the class on an instance, and a random variable vector X which represents the observed values of attributes. Assume c a label of a determined class and x a test instance x to be classified. The most probable class will be the one of most high value to $P(C = c|X = x)$, which is the probability of occurrence of c class given a instance x. Equation (1) represents the Bayes' rule applied to calculate this probability [15].

$$p(C = c|X = x) = \frac{p(C = c)p(X = x|C = c)}{p(X = x)} \quad (1)$$

This algorithm creates a probability table where it learns to determine the class of an attribute from the database, *i.e.*, if a new unknown record is submitted to the table, the algorithm returns the probability to which class the new record belongs. Table V shows the form that the preprocessed sentence resembles, which is ready to start training by the algorithm.

Using the NLTK tool, the table of probabilities that the naive Bayes classifier uses for training and learning new attributes is built. This table is built with the following command.

TABLE V
EXAMPLE OF PHRASES CLASSIFIED AS HAPPINESS OR SADNESS BY THE NAIVE BAYES ALGORITHM.

Original phrase	Preprocessed phrase	Class
My life is great, I'm happy	lif great happ	Happiness
I'm feeling happy	feel happ	Happiness
I'm scared of getting sick	scar get sick	Sadness
it makes me afraid	mak afraid	Sadness

```
classifier =
nltk.NaiveBayesClassifier.train(comptrainingbase)
```

Then, the algorithm already has information about the probability of each word in relation to its particular class and is capable to predict new unknown cases to a class. Therefore, the algorithm can already classify a sentence. For this to occur, is necessary to perform the following command.

```
print(classifier.classify(new))
```

This is the command in which the texts are classified as happiness or sadness, using the naive Bayes algorithm and the `classify` function. The classified text with its respective emotion is determined through probabilities present in the naive Bayes classifier. Still using NLTK tools, the probability of the class that was predicted to be correct can be generated.

For example, the algorithm receives the phrase "I'm fine, I just have to thank for this wonderful life" as a parameter and returns as an answer that the phrase can belong to the happiness class, with a probability of 99.55%, or to the sadness class, with a probability of 0.45%.

IV. PERFORMANCE EVALUATION AND RESULT ANALYSIS

The predictive model evaluation is based on common metrics used in the literature. These metrics show successes and fails for a given data set. The metrics used for the model assessment tools are detailed below.

In the ML field, and specifically in classification, the confusion matrix is used as a way of visualizing the performance of a certain algorithm, verifying that the predictive model is not confusing any class.

In this research, the classification problem consists of identifying whether a new text belongs to the happiness or sadness classes correctly. The predicted results must be compared with the real cases, thus providing the true positive (TP), true negative (TN), false negative (FN), and false positive (FP) indicators. These metrics are used to calculate the metrics associated with the confusion matrix.

Four metrics are used to analyze the performance, to know, precision, recall, F-score, and accuracy. Below are each of these metrics.

Accuracy is the measure used to analyze the classification performance of records considered as positive. In this case, the greater the precision, the greater the TP exactness.

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

The recall analyzes the TP rate, *i.e.*, the number of cases that the classifier correctly predicts the TPs, as well as in the precision, however, using the FN indicator. A low percentage on this metric indicates that the algorithm classifies many positive records as FN.

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

The F-Score represents the harmonic average of the two metrics presented previously to achieve at a number that indicates the general quality of the model. This indicator works well even with databases that have disproportionate classes. With this measure, it is possible to assess the classifier performance through just one indicator, using the harmonic average.

$$F - Score = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \tag{4}$$

Accuracy represents the correctness rate for the entire model. The best result for an algorithm, concerning the confusion matrix indicators, is to present 100% accuracy. In this case, only the main diagonal of the matrix would be filled.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{5}$$

According to Vanbelle, the Kappa coefficient, proposed by Cohen in 1960, is a robust weighted measure that takes into account the hits and misses according to the confusion matrix [16]. This coefficient is expressed by Equation (6).

$$K = \frac{P_0 - Pe}{1 - Pe} \tag{6}$$

Where

$$P_0 = \frac{\sum_{i=1}^{M}(n_d)}{N} \tag{7}$$

$$P_e = \frac{\sum_{i=1}^{M}(n_{lc})}{N^2} \tag{8}$$

n_d is the total number of the main diagonal of the confusion matrix; n_l is total for row i of the matrix and n_c is total for column j; M is the total number of classes, and N is the total number of decisions in the matrix. This coefficient is commonly used for training evaluation terms. From it, it is possible to distinguish in which situation the agreement of the evaluations is concerning a reference system. Table VI shows the agreement degree about the Kappa coefficient.

TABLE VI
KAPPA COEFFICIENT INTERPRETATION.

Kappa Coefficient	Concordance degree
< 0.00	very poor
0.00 - 0.20	poor
0.21 - 0.40	weak
0.41 - 0.60	regular
0.61 - 0.80	good
0.81 - 0.99	excellent
1.00	perfect

The confusion matrix was built using NLTK tools. With the matrix obtained, it was possible to calculate the performance indexes of the model using the metrics discussed in the previous session. Table VII shows the model performance based on the metrics of the confusion matrix. The results are presented in the form of the value obtained by the formulas and their percentage value. The results were considered satisfactory for these performance measures.

TABLE VII
MODEL PERFORMANCE INDEX.

Metrics	Value	%
Precision	0.8392	83.92
Recall	0.9400	94.00
F-Score	0.8866	88.66
Accuracy	0.8636	88.36

Table VIII describes the model performance concerning the Kappa coefficient. A result of 0.8626 was obtained, as can be seen highlighted. The model achieved a performance with an "excellent" agreement degree.

TABLE VIII
KAPPA INDEX OF THE MODEL.

Kappa coefficient	Obtained coefficient	Concordance degree
< 0.00	-	very poor
0.00 - 0.20	-	poor
0.21 - 0.40	-	weak
0.41 - 0.60	-	regular
0.61 - 0.80	-	good
0.81 - 0.99	**0.8628**	**excellent**
1.00	-	perfect

V. CONCLUSION

Text mining is defined as an ML area where the main objective is to develop methods to explore a set of textual data and extract important information. In this study, text mining and ML techniques in textual data of pregnant women are explored to prevent gestational depression.

Through the development of this research, it was possible to validate a model for emotion classification in texts, a technique in which its use in medicine tends to increase. The goal is to provide for healthcare professionals a leading tool to aid gestational depression prevention, thus avoiding more complex risk situations.

Excellent results were obtained through the metrics associated with the confusion matrix, analyzing the model performance results. The following results were obtained. Precision (83.92%), recall (94.00%), F-Score (88.66%), accuracy (86.36%), and Kappa coefficient (0.8628). The results validate the use of the model for the analysis of feelings in new texts when submitted to the analysis of the database of this study. The performance of the proposed solution concerning new cases was satisfactory, classifying 88,36% of data correctly.

As further work, it is expected to add other emotions and increase the database, improving the predictive model using novel ML techniques. Besides, to develop a Web application to make the solution a tool for daily use by pregnant women and health professionals for better monitoring the high-risk pregnancy.

ACKNOWLEDGMENTS

This work is supported by FCT/MCTES through national funds and when applicable co-funded EU funds under the Project UIDB/50008/2020; and by Brazilian National Council for Scientific and Technological Development - CNPq, via Grant No. 309335/2017-5.

REFERENCES

[1] J. Potter and A. Hepburn, "Shaming interrogatives: Admonishments, the social psychology of emotion, and discursive practices of behaviour modification in family mealtimes," *Br. J. Soc. Psychol.*, vol. 59, no. 2, pp. 347–364, 2020.

[2] J. A. Horowitz, B. Posmontier, L. A. Chiarello, and P. A. Geller, "Introducing mother-baby interaction therapy for mothers with postpartum depression and their infants," *Arch. Psychiat. Nurs.*, vol. 33, no. 3, pp. 225–231, 2019.

[3] F. Greco and A. Polli, "Emotional text mining: Customer profiling in brand management," *Int. J. Inf. Manage.*, vol. 51, p. 101934, 2020.

[4] E. Cambria, "Affective computing and sentiment analysis," *IEEE Intell. Syst.*, vol. 31, no. 2, pp. 102–107, 2016.

[5] E. Reategui, M. Klemann, and M. D. Finco, "Using a text mining tool to support text summarization," in *12th International Conference on Advanced Learning Technologies (ICALT)*, (Rome, Italy), pp. 607–609, IEEE, Jul. 4-6 2012.

[6] E. Saravia, C. Argueta, and Y.-S. Chen, "Emoviz: Mining the world's interest through emotion analysis," in *International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, (Paris, France), pp. 753–756, IEEE/ACM, Aug. 25-28 2015.

[7] R. S. Muniz, R. G. Rodrigues, and G. P. Guedes, "Croca - cromoterapia e computação afetiva: auxiliando os estados de ansiedade," in *XXII Simpósio Brasileiro de Sistemas Multimídia e Web (WebMedia)*, (Teresina, PI, Brazil), pp. 145–148, SBC, Nov. 8-11 2016.

[8] N. P. Brito, "Puerpério APP-uma aplicação móvel de apoio à puerpéras com depressão pós-parto baseada na escala de edimburgo," Master's thesis, Universidade Federal do Amazonas (UFAM), Itacoatiara, AM, Brazil, 2019.

[9] N. Warikoo, Y.-C. Chang, H.-J. Dai, and W.-L. Hsu, "An ensemble neural network model for benefiting pregnancy health stats from mining social media," in *Asia Information Retrieval Symposium*, (Taipei, Taiwan), pp. 3–15, Springer, Nov. 28-30 2018.

[10] Y. Chen, B. Zhou, W. Zhang, *et al.*, "Sentiment analysis based on deep learning and its application in screening for perinatal depression," in *IEEE Third International Conference on Data Science in Cyberspace (DSC)*, (Guangdong, China), pp. 451–456, IEEE, Aug. 18-21 2018.

[11] F. L. Seixas, C. M. Carvalho, D. C. Muchaluat-Saade, A. Conci, and J. Laks, "Descrição de um sistema de suporte ao diagnóstico de demência e transtornos mentais relacionados," *J. Health Inform.*, vol. 9, no. 3, 2017.

[12] N. Hardeniya, J. Perkins, D. Chopra, *et al.*, *Natural language processing: Python and NLTK*. Birmingham, UK: Packt Publishing, 2016.

[13] A. V. Kunte and S. Panicker, "Using textual data for personality prediction: a machine learning approach," in *4th International Conference on Information Systems and Computer Networks (ISCON)*, (Mathura, Uttar Pradesh, India), pp. 529–533, IEEE, Nov. 21-22 2019.

[14] A. Hasan, S. Moin, A. Karim, and S. Shamshirband, "Machine learning-based sentiment analysis for Twitter accounts," *Math. Comput. Appl.*, vol. 23, no. 1, p. 11, 2018.

[15] A. R. Webb, *Statistical pattern recognition*. John Wiley & Sons, 2003.

[16] S. Vanbelle, "A new interpretation of the weighted kappa coefficients," *Psychometrika*, vol. 81, no. 2, pp. 399–410, 2016.