Received: 18 November 2018 | Revised: 20 February 2019 | Accepted: 15 March 2019
DOI: 10.1111/exsy.12409

WILEY Expert Systems

ORIGINAL ARTICLE

Prediction of postpartum depression using machine learning techniques from social media text

Iram Fatima[1] | Burhan Ud Din Abbasi[2] | Sharifullah Khan[2] | Majed Al-Saeed[1] | Hafiz Farooq Ahmad[1] | Rafia Mumtaz[2]

[1]College of Computer Sciences and Information Technology, King Faisal University, Hofuf, Saudi Arabia
[2]School of Electrical Engineering and Computer Science, National University of Sciences and Technology, Islamabad, Pakistan

Correspondence
Iram Fatima, College of Computer Sciences and Information Technology, King Faisal University, Hofuf, Saudi Arabia.
Email: ialrehman@kfu.edu.sa

Funding information
Deanship of Scientific Research, King Faisal University, Grant/Award Number: 180055

Abstract

Early screening of mental disorders plays a crucial role in diagnosis and treatment. This study explores how data-driven methods can leverage the information available on social media platforms to predict postpartum depression (PPD). A generalized approach is proposed where linguistic features are extracted from user-generated textual posts on social media and categorized as general, depressive, and PPD representative using multiple machine learning techniques. We find that techniques used in our study exhibit strong predictive capabilities for PPD content. Holdout validation showed that multilayer perceptron outperformed other techniques such as support vector machine and logistic regression used in this study with 91.7% accuracy for depressive content identification and up to 86.9% accuracy for PPD content prediction. This work adopts a hierarchical approach to predict PPD. Therefore, the reported PPD accuracy represents the performance of the model to correctly classify PPD content from non-PPD depressive content.

KEYWORDS

machine learning, mental health, moods and emotions, postpartum depression, social media

1 | INTRODUCTION

Transition to parenthood is one of the major phases in lives of people impacting various aspects of life, at times even causing negative emotional impact (Hudson, Elek, & Campbell-Grossman, 2000). These changes seem to affect mothers and fathers both because of their inability to resolve differences between personal, social, and professional lives (Genesoni & Tallandini, 2009; Woolhouse, McDonald, & Brown, 2012). Postpartum depression (PPD) is one of the more common disorders diagnosed in parents. *Diagnostic and Statistical Manual of Mental Disorders* defines PPD as a major depressive disorder with peripartum onset with the most recent episode occurring from anywhere during pregnancy till 4 weeks after childbirth (American Psychiatric Association, 2013). *International Classification of Diseases* recognizes this disorder up to the period of 6 weeks after childbirth (World Health Organization, 2004). Percentage of individuals affected from this disorder shows large variation around the world and can be as high as 63% (Kalyani, Saeed, Rehman, & Mubbashar, 2001). Although mothers are more susceptible to PPD, an estimated 4% of fathers also experience this disorder (Davé, Petersen, Sherr, & Nazareth, 2010). A study found that 8% of adoptive mothers also experienced depression possibly due to lifestyle changes (Mott, Schiller, Richards, O'Hara, & Stuart, 2011). On average, 15% of mothers are expected to be suffering from PPD all over the world. Because no biological measure has been identified to be the cause of PPD, it becomes a challenge to diagnose PPD considering that changes in appetite, sleep patterns, and excessive fatigue are a norm for women after childbirth (Pearlstein, Howard, Salisbury, & Zlotnick, 2009).

Researchers have identified various factors as predictors of PPD in individuals who are suffering or at risk (Beck, 1998, 1998, 2001; Reck, Stehle, Reinig, & Mundt, 2009). Some researchers have designed a series of questions to be answered in order to simplify the process of identification of PPD (Cox, Holden, & Sagovsky, 1987). Similarly, some have worked on measuring the severity of the depression (Kroenke, Spitzer,

Abbreviations: D-CC, depressive content classification; PPD, postpartum depression; PPD-CC, postpartum depression content classification.

& Williams, 2001). These methods, although useful, require individuals to explicitly answer a series of questions for the sake of diagnosis. Since the presence of a mental health professional is critical for evaluation of answers, this approach remains limited to the point of availability of professionals in a region and general awareness among public regarding importance of mental health (Mohr et al., 2010). Unavailability of mental health professionals and high costs associated with treatment is a challenge. Hence, the need exists for low-cost and innovative methodologies for identification of individuals suffering from PPD and/or detecting tendencies of developing PPD.

On the other side, communication channels have evolved a great deal due to the technology-driven innovation in last few decades. Social media platforms like Twitter[1], Facebook,[2] and reddit[3] have become go to places for expressing opinions and networking with like-minded people (Hanna, Rohm, & Crittenden, 2011; De Choudhury & De, 2014). Increase in the number of online parenting social platforms indicates that there are significant number of users seeking advice on health for the purpose of self-diagnosis (Sarkadi & Bremberg, 2005). As a consequence of mental health conditions, people show change in usage patterns as well as the nature of content they share on social media (De Choudhury, Counts, & Horvitz, 2013; De Choudhury & De, 2014). Machine learning techniques give computer systems the ability to learn from the data, without the need of being explicitly programmed for a certain task. Online social platforms can provide enough information to enable reliable prediction scores using machine learning, thus allowing relevant authorities and organizations to plan for early intervention mechanism to prevent and handle PPD cases in an effective and timely manner.

We note that there is little prior research on prediction of PPD using machine learning on social media text. Moreover, existing work is focused towards solutions that employ platform specific features, such as interactivity with friends, post likes, and others to study and predict users experiencing PPD (De Choudhury, Counts, Horvitz, & Hoff, 2014). The approach is limited to a specific platform and hard to generalize for other platforms. In this research, we consider using linguistic features to propose a solution that can be generalized and deployed across online social platforms. The linguistic features are extracted using linguistic inquiry word count (LIWC). Moreover, the proposed approach distinguishes PPD from non-PPD depressive posts. This approach is two layered: In the first layer, it categorizes between general and depressive discussions, and in the second layer, it distinguishes PPD discussions from non-PPD depressive posts. Using layered approach, we show that not only depressive content (PPD + non-PPD) can be predicted from normal discussion but PPD and non-PPD content can also be bifurcated using the same feature set.

In the section that follows, we present literature review for the reader in order to build a sense of this domain. The section that comes after contains details on the proposed methodology. We discuss in detail following steps: data collection, feature selection, and prediction techniques. Evaluation measures and experimental results section is followed by the conclusion and discussion for future work in the last section.

2 | RELATED WORK

Research by Rude, Gortner, and Pennebaker (2004) found a significant difference in language of currently depressed individuals when compared with that of never-depressed individuals. Attempts to identify risk factors for PPD have been ongoing for decades, resulting in a wide variety in nature of factors ranging from mistrust and marital problems to past history of PPD (Braverman & Roux, 1978; Boyer, 1990). Although survey-based methods are not alternatives for clinical judgement, questionnaires have been developed to aid in screening process (Cox et al., 1987). In recent years, there have been effort to measure the impact of mindfulness-based cognitive therapy in PPD-suffering mothers and for prevention of its recurrence in pregnant women (Shulman et al., 2018; Dimidjian et al., 2015).

From technological viewpoint, researchers are also focusing towards provision of systems that can accomplish the task of screening and generate alerts based on behavioural attributes as presented by data on social networking platforms. Research found that web-based interventions showed promising results over a wide range of mental health issues and recommended focus on improvements in data collection and analytical systems (Mohr, Burns, Schueller, Clarke, & Klinkman, 2013). Another research concluded that non-professional support available through online platforms was helpful for individuals facing emotional issues (Baumel, 2015). A system based on ensemble machine learning techniques to classify individuals at risk of major depressive disorders from their Facebook activities was proposed (Hussain et al., 2015). They used the number of friends, followers, status updates, and interactivity based on comments and likes to measure changes in routine, help seeking, and drug references. Researchers have also identified a set of features that can predict depressive posts and communities, additionally predicting severity of depressive posts based on mood tags as available on source platform (Fatima, Mukhtar, Ahmad, & Rajpoot, 2017). Although such tools can aid in screening and early intervention, their scope remains limited to a certain platform due to the use of platform specific features. Moreover, machine learning techniques used in these studies are known to be either limited in their learning capabilities or suffer from high computational costs in the presence of large training datasets.

A recent research found higher use of absolutist words in mental health-related forums (Al-Mosaiwi & Johnstone, 2018). Their work was based on custom dictionaries of 19 absolutist and 43 nonabsolutist words in addition to features extracted through LIWC tool. Their experiments were focused on major mental disorders like depression, bipolar disorder, and post-traumatic stress disorder.

[1]https://twitter.com/.
[2]https://www.facebook.com/.
[3]https://www.reddit.com/.

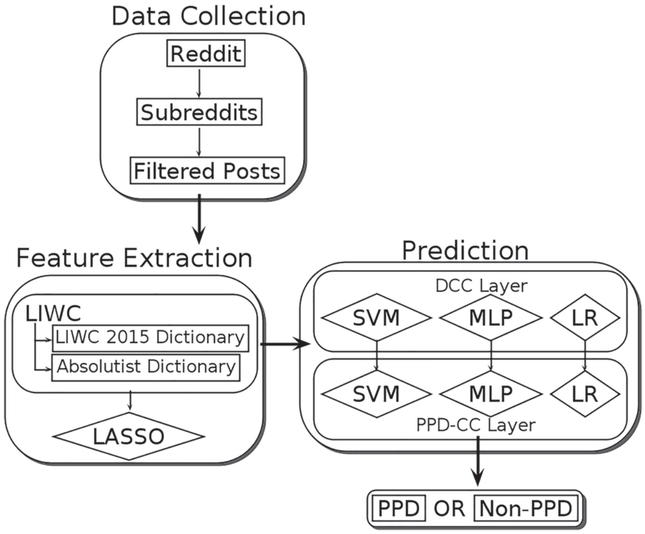


FIGURE 1 Process model of proposed layered approach to postpartum depression (PPD) prediction. LIWC, linguistic inquiry word count; LR, logistic regression; LASSO, least absolute shrinkage and selection operator; MLP, multilayer perceptron; PPD-CC, PPD content classification; SVM, support vector machine

In another research (De Choudhury, Counts, & Horvitz, 2013), social media-based behavioural markers were used to predict posts indicative of depressive tendencies, showing the possibility of large-scale adoption of such system to monitor mental health issues in large populations. They used crowd sourcing to collect data and identify individuals who had been clinically diagnosed as depressed and their questionnaire-based assessment also indicated presence of depression. In a different study, researchers (De Choudhury et al., 2014) predicted PPD suffering individuals based on social media user profile using a number of platform-related features, such as status updates, comments, and wall posts; along with linguistic features such as type of pronouns; and some personal and demographic data such as income, ethnicity, and occupation. It was found that mothers using Facebook preferred not to disclose real mental state and feelings related to PPD due to the reason that Facebook friends are usually our physical world contacts too. In addition to the variation in privacy settings of Facebook users, there is the limitation of availability of personal information such as income that renders methodology adopted by user-specific studies less scalable.

3 | METHODOLOGY

The scope of this study is to propose an approach for prediction of PPD using machine learning on the natural language data shared on social media. We speculate that the linguistic patterns as manifested by social media text should provide enough information to enable reliable prediction scores, thus allowing relevant authorities to formulate early intervention mechanisms for prevention and treatment of PPD cases in a cost-effective and timely manner. We try to address the challenges as observed in literature review to find a generalized approach that is not platform dependent and explore whether absolutist words can further be used by machine learning techniques as an indicator of PPD. We use text present in the body of the post for this purpose. We refers to it as "post" or "content" and use both terms interchangeably.

We explore the use of textual posts on social media to identify PPD-related discussions. Social media sites give various privacy options to their users, thus providing controls regarding information flow to their peers. Generally, user profiles do not contain medical history or diagnosis. It was observed that at times, users also post in PPD groups to seek help for their loved ones. However, the content of posts usually revolves around the challenges faced by the individual suffering from PPD. We use machine learning techniques to classify general discussion, PPD, and non-PPD depressive content based on the linguistic features only. Because PPD is recognized as a major depressive disorder indicating the presence of a taxonomic scheme, we adopt a layered approach that first bifurcates depressive content and general discussions, and at the second stage, PPD content is identified from depressive content as shown in Figure 1. The proposed approach is explained in following subsections.

3.1 | Data collection

In this study, our approach and choice for the data source was driven by the findings regarding differences observed in online activity of mothers. Moreover, by focusing on linguistic patterns, we try to address shortcomings of the previous works that focused on user profiles, individual activity patterns, and platform-specific features. In this research, we selected reddit, a widely used social media platform that allows users to post and interact with others anonymously. In other words, it allows users to maintain a layer of privacy and share their problems candidly; this as found by De Choudhury et al. (2014), can be a factor stopping PPD-affected mothers from opening up about their mental state on social media platforms. Members, also known as "redditors," post links and text in communities referred to as "subreddits." Users can comment on the posts and comments, forming trees of comments, thus enabling users to engage with each other at multiple levels. Because reddit is based on communities, people with similar interests carry out discussions on subreddits related to their areas of interest. Strong content moderation culture of reddit usually ensures that off-topic conversations are removed and frequent violators are banned. We used PRAW (Python Reddit API Wrapper) to collect data from reddit including posts and associated metadata from several subreddits. Posts from 21 subreddits were collected of duration starting from January 2011 till April 2018. Because the goal of the study was to explore the possibility of prediction based on linguistic features, during preprocessing steps, all such posts that only had a title and did not contain text in the body of post were removed. Similarly, all image, video, and link-based posts were filtered out. Each post can be associated to a general category based on the nature of groups, that is, general discussion, depression-related, and PPD posts. Data for each group were gathered from a number of subreddits with the purpose of having diverse and well-represented content. For the purpose of clarity, we assign each group a specific name, such as, daily life group, depression group, and PPD group. Details on each group are given in Table 1.

It is worth noting that PPD group included dedicated PPD subreddits as well as other subreddits that though not primarily focused towards PPD fall under the topics closely related to parenthood such as parenting, breastfeeding, and baby bumps. These subreddits were included because of the low number of posts in PPD-focused subreddits. For non-PPD-focused subreddits, we only considered posts that contained the term "postpartum depression" in the title or the body of post. Terms like PPD, antenatal depression, and perinatal depression were not considered for the creation of this dataset. Size of dataset was restricted due to unavailability of large number of PPD posts. Due to relatively less number of PPD-related posts, we randomly selected 800 posts from the data crawled from *r/depression* and *r/depression_help*; similarly, 1,588 posts were randomly selected from the collection of posts in *Daily Life Group*. The posts selected from depression and PPD group were 800 and 788, respectively, which make them 1,588. The similar number, that is, 1,588, were selected from "not depressive" class to balance the representation of classes in the dataset.

3.2 | Feature extraction

This work focused on using the content of text-based posts and categorically avoided the use of metadata such as upvotes, downvotes, score, number of comments, date, and time of posts. For the purpose of extracting linguistic features, a widely used resource LIWC tool was used (Pennebaker, Francis, & Booth, 2001). LIWC analyses text by comparing and calculating percentage of words that match built-in dictionaries, such as LIWC-2001, LIWC-2007, and LIWC-2015. LIWC-2015 dictionary gives 93 features for the each post ranging from measures like word count, words per sentence, and emotional tone to first-person singulars, interrogatives, and comparisons. In the light of findings by Al-Mosaiwi and Johnstone (2018), we used LIWC to calculate values based on custom dictionary of absolutist words, increasing the number of linguistic features to 94 for each post in dataset. Absolutist dictionary composed of 19 words, which are as follows: *absolutely*, *all*, *always*, *complete*, *completely*, *constant*, *constantly*, *definitely*, *entire*, *ever*, *every*, *everyone*, *everything*, *full*, *must*, *never*, *nothing*, *totally*, and *whole*. Table 2 contains some examples for postcontent and their respective feature values.

Features used for processing natural language usually consist of words, phrases, or their numeric representations to fit statistical models of linguistic concepts. Using swarm plots for features, such as "emotional tone" and "analytical thinking", as shown in Figure 2, helped us to obtain

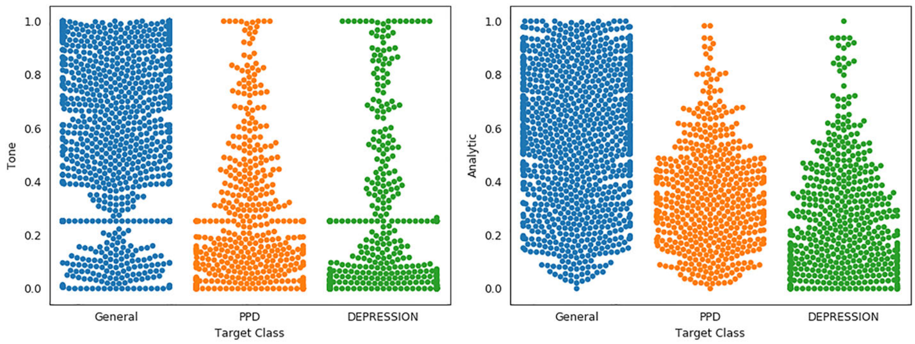
TABLE 1 Groups and list of included subreddits

Group	Posts	Subreddit
Daily life group	1,588	r/books, r/business, r/movies, r/technology, r/graphic_design
Depression group	800	r/depression, r/depression_help
PPD group	788	r/postpartumdepression, r/MyPPDSupport, r/Parenting, r/relationships, r/childfree, r/beyondthebump, r/mommit, r/BabyBumps, r/AskWomen, r/AskReddit, r/legaladvice, r/JUSTNOMIL, r/Breastfeeding, r/breakingmom

Abbreviation: PPD, postpartum depression.

TABLE 2 Examples of post content and their normalized feature values

Post content	Health	Work	Tone	Affiliation
Fear, pain, happiness, sadness, all non existent. Something I wish was possible	0.69	0	0	0
Looks like a fun exercise. Lots of really good concepts on the Dribbble and Instagram!	0.55	0	1	0.33
Just had my first baby a week ago, but my girlfriend is pretty sad. Any suggestions on how to cheer her up? Im planning on taking her out for a picnic tomorrow. Any suggestions?	0	0	0.80	0.14
Looking for a site which had a few different templates on how to reply to clients asking for work with different budgets etc. Will delete this and maybe post when I get the answer.	0	0.27	0.25	0



(a) Emotional Tone (b) Analytical Thinking

FIGURE 2 Distribution of data points for target classes. PPD, postpartum depression

an idea of underlying distribution of data points for each class. Tone in Figure 2a represents the extent of "emotional thinking" by individuals of each group. Similarly, analytic in Figure 2b is representative of analytical thinking as presented by linguistic patterns of each group. From the set of 94 features as representative measures, we were interested in identifying a list of most prominent features in terms of their ability to describe a response variable. This would allow for easier interpretation of the model, faster performance by the algorithms even for a large dataset ,and in reducing overfitting of the model. To achieve this goal, we used least absolute shrinkage and selection operator (LASSO; Tibshirani, 1996). LASSO works by employing ridge regression, which is the process of continuous shrinking coefficient values for a variable and setting some coefficients to zero. This becomes particularly important for our purpose because of the presence of near-linear relationship between independent features (collinearity), for example, word count, first-person singulars, and comparisons. Because LASSO uses the knowledge of target class, the target class given for the purpose of feature selection was based on previously described groups, such as daily life group was marked as *nondepressive*, depression group was marked *depressive*, and PPD group was marked *PPD class*. The optimization objective for LASSO, as defined in scikit-learn documentation is

$$(1/(2 * n\_samples)) * ||y - Xw||^2 + \text{alpha} * ||w||, \tag{1}$$

where y is the predicted outcome using prediction function Xw in Equation (1), term $||y - Xw||^2$ represents sum of square errors, $||w||$ is the ℓ_1−norm of parameter vector, and alpha is a constant (Pedregosa et al., 2011).

LASSO uses a tunning parameter as a measure of the penalty. As its value increases, the coefficients are forced to become zero. During selection process, only features having nonzero coefficients are selected. For the value 0.0035 of alpha, LASSO model returned 20 nonzero coefficient variables. Features associated with nonzero coefficients were selected to serve as a feature set for the task of prediction. Details on the feature set are given in Table 3.

The feature set represents multiple aspects of life, such as concerns related to family and home, mentions of emotional states, discussion of goals, and feelings as expressed by people in their posts. Summary variables of LIWC like word count, emotional tone, and dictionary words were found to be important predictors of target classes.

3.3 | Prediction

As discussed in Section 3.1, the dataset composes of three groups, therefore representing variety in the nature of posts from the perspective of general content, depressive, and PPD posts. In our design, there are two layers: first layer referred to as depressive content classification (D-CC) and second layer referred to as PPD content classification (PPD-CC). D-CC layer classifies between general posts and depressive posts, whereas PPD-CC layer classifies between PPD and non-PPD posts as shown in Table 4 . We used three different machine learning techniques for prediction on the dataset. These techniques included logistic regression (LR), support vector machine (SVM), and multilayer perceptrons (MLPs). In each of the three techniques, that is, SVM, LR, and MLPs, we trained two models, one for each layer in hierarchical classification such as D-CC and PPD-CC layers. Each model was trained using same feature set as selected by LASSO.

The goal of D-CC layer is to train a machine learning model that can learn from the feature set, the distinguishing patterns, which can be used to identify depressive content from daily life discussions relating to professions, entertainment, and hobbies. We trained models for each classifier using the feature set extracted by LASSO and compared the performance of each classifier. Table 4 shows target classes assigned to each posts in the dataset with respect to each layer. The flow of whole classification process is shown in Algorithm 1. Class labels are assigned to each post from lines 2–11. Feature extraction based on LIWC-2015 and absolutist dictionary is given in lines 12 and 13. Data preprocessing, feature selection, and model training steps have been listed from lines 15–21. Line 22 and onwards list the steps of model validation phase for the proposed layered approach.

Content, representative of PPD, is classified in this step as PPD, thus separating it from non-PPD depressive content as predicted in D-CC layer. The output of D-CC layer, that is, depressive content is introduced in PPD-CC layer for the task of identifying PPD posts from non-PPD

TABLE 3 Selected feature set based on LASSO and their detail based on LIWC

Category	Feature	Examples
Social words	Family	daughter, dad, aunt
Absolutist words	Absolutist	always, must, entire
Function words	Negations	no, not, never
	1st-person singular	I, me, mine
	1st-person plural	we, us, our
	3rd-person singular	she, her, him
	Impersonal pronouns	it, it's, those
Personal concerns	Death	burry, coffin, kill
	Home	landlord, kitchen
Drives and needs	Affiliation	ally, friend, social
	Drive	ego, purpose, ambition
Time orientations	Focus present	today, is, now
Psychological processes	Anger	hate, kill, annoyed
	Negative emotion	hurt, ugly, nasty
Perceptual processes	Feeling	feels, touch
Other grammar	Interrogatives	how, when, what
	Comparatives	greater, best, after
Word count	WC	—
Summary variables	Dictionary words	—
	Emotional tone	—

Abbreviations: LASSO, least absolute shrinkage and selection operator; LIWC, linguistic inquiry word count.

TABLE 4 Target classes assigned to groups for hierarchical training

Group	D-CC	PPD-CC
Daily life group	Not depressive	Not PPD
Depression group	Depressive	Not PPD
PPD group	Depressive	PPD

Abbreviations: D-CC, depressive content classification; PPD, postpartum depression; PPD-CC, postpartum depression content classification.

depressive posts. As given in Table 4, posts in daily life group were also tagged as *not PPD* in order to handle false positives in the output of D-CC layer, thus enabling the classifier to handle any general discussion posts incorrectly classified as depressive in D-CC layer.

3.3.1 | Logistic regression

LR is a linear model that performs prediction by considering input values of the feature set and a bias term in order to get a model that fits the training set. A regularization term is introduced, which prevents models from overfitting, thus giving us control over the complexity of model. This complexity is quantified using L2 regularization that uses least squares error, minimizing the sum of the square of difference between estimated and target values (Tikhonov, 1977). For the choice of algorithm to be used for optimization objective, we use Newton's method. Although computationally expensive, it performed reasonably for our dataset (Fletcher, 1987).

3.3.2 | Support vector machines

It is a widely used technique that aims to find a boundary (hyperplane) that divides data into two classes such that there exists greatest possible distance between points of training set and the hyperplane (Schlkopf & Smola, 2002). For cases when dataset is complex and there is no clear hyperplane, data are mapped into higher dimensions in order to find a hyperplane that can segregate the data, known as kernelling. Because the dataset was not huge, we used radial basis function kernel. Radial basis function kernel maps non-linear data into higher dimensions, therefore handling the case of non-linear relation between features and target class. Moreover, misclassification penalty parameter was set to 1, thus allowing the model to generalize and not over fit the training data (Hsu, Chang, & Lin, 2003; Scholkopf et al., 1997).

Algorithm 1 Pseudocode for layered depressive and PPD content classification

```
INPUT:
  Data set DS
  Daily Life Group DLG[]                         ▷ Lists of communities/subreddits
  Depressive Group DG[]                          ▷ Lists of communities/subreddits
  PPD Group PPDG[]                               ▷ Lists of communities/subreddits
OUTPUT:
  Depressive Content Layer D_CC[]                ▷ Layer 1 Output
  PPD Content Layer PPD_CC[]                     ▷ Layer 2 Output
    Algorithm:
for i ← 1 : length(DS) do
    if DS[i]['source_subreddit']inDLG[] then
        DS[i]['Class_DCC'] ←' nonDepressive'
        DS[i]['Class_PPD'] ←' nonPPD'
    else if DS[i]['source_subreddit']inDG[] then
        DS[i]['Class_DCC'] ←' Depressive'
        DS[i]['Class_PPD'] ←' nonPPD'
    else if DS[i]['source_subreddit']inPPDG[] then
        DS[i]['Class_DCC'] ←' Depressive'
        DS[i]['Class_PPD'] ←' PPD'
    end if
    DS ← LIWC_2015(DS[i]['post_content'])
    DS ← LIWC_Absolutist(DS[i]['post_content'])
end for
Test, Train ← Test_Train_Split(DS)
Train[AllFeatures[]] ← Normalize(Train[AllFeatures[]])
coefficient_value[], feature_name[] ← LASSO(Train[AllFeatures[]], Train['Class_DCC'])
FeatureSet[] ← NonZeroCoEfficients(coefficient_value[], feature_name[])
classifier ← Multi_Layer_Perceptron(length(FeatureSet[]), 20, 20, 2)
DCC_Layer_Model ← classifier.Train(Train[FeatureSet[]], Train['Class_DCC'])
PPD_Layer_Model ← classifier.Train(Train[FeatureSet[]], Train['Class_PPD'])
Test[AllFeatures[]] ← Normalize(Test[AllFeatures[]])
D_CC[] ← DCC_Layer_Model.Predict(Test[FeatureSet[]])
predicted_depressive ← Filter_Posts(whereD_CC[]is'Depressive')
PPD_Test ← Filter_Posts(whereTestinpredicted_depressive)
PPD_CC[] ← PPD_Layer_Model.Predict(PPD_Test[FeatureSet[]])
```

3.3.3 | Multilayer perceptron neural networks

MLPs are feedforward neural networks that employ standard back propagation algorithm (Gardner & Dorling, 1998). A number of layers of neurons work together to learn from the dataset. The first layer takes the data input, whereas the last layer produces the final output. The layer between input and output layers is also known as hidden layers. Hidden layers can contain any number of neurons, therefore increasing the ability of the whole network to learn from the data. Similar to other machine learning algorithms, a balance in parameter values is needed for optimal performance. In MLP, a balance between the number of layers and neurons should be maintained. A very low number of layers or neurons can hinder the ability of network to learn efficiently, whereas a very high number can increase the computational cost significantly and also make the network learn unnecessary details of a training set, therefore causing the model to overfit, thus rendering it less useful for unseen data. Each connection between neurons is assigned a random weight initially and later updated based on the learning of network as seen by examples in the training set. Each neuron processes the input based on a predefined function referred to as activation function. We used four-layered MLP, an input layer with neurons equal to length of feature set, two hidden layers having 20 neurons in each layer, and an output layer composing of two neurons. Rectified linear unit function was used as activation function with stochastic gradient-descent-based optimizer and using constant learning rate of 0.0001.

4 | RESULTS AND DISCUSSION

This section explains briefly characteristics of dataset, measures used to evaluate the performance of proposed methodology, results, and comparisons of techniques and discussion on results.

4.1 | Dataset description

As shown in Table 4 for D-CC layer, depression group and PPD group were assigned target class as "depressive," making both classes of D-CC layer (depressive/nondepressive) well represented, thus enabling machine learning models to learn equally well regarding features of each class. In the dataset composed of 3,176 posts, approximately 50% were from daily life group hence tagged as nondepressive in nature, 25% of posts belonged to depression group, and remaining 25% were from the PPD group . In order to evaluate the performance of our models, 67.33% split of posts was adopted, yielding a training set of 2,127 posts and a testing set or holdout set of 1,049 posts. Training was performed through 10-fold crossvalidation (C.V.) scheme, where each iteration used ninefolds of the training set to train the model and onefold for validation purpose.

4.2 | Evaluation measures

Accuracy was used as the primary measure to gauge the performance of each model. To evaluate the reliability of our results, we report for each model macrolevel precision and recall scores. Accuracy, precision, and recall are defined in Equations 2, 3, and 4, respectively.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}, \tag{2}$$

$$\text{Precision} = \frac{TP}{TP + FP}, \tag{3}$$

$$\text{Recall} = \frac{TP}{TP + FN}, \tag{4}$$

where *TP*, *TN*, *FP*, and *FN* are the number of true positives, true negatives, false positives, and false negatives, respectively (Manning, Raghavan, & Schütze, 2008). Macroprecision is mean of precision scores calculated separately for each target label, whereas macrorecall is the mean of recall scores computed individually for each label. Scores for each performance measure are reported for both 10-fold C.V. as well as holdout validation.

Tenfold C.V. was used to evaluate and optimize training results. Stratified K-fold scheme was used for C.V. in order to ensure that all folds maintain percentage representation for each class (Kohavi, 1995). Moreover, in order to test the prediction capability of our model on unseen data, we select the best performing model in 10-fold C.V. and perform holdout validation on holdout set. We discuss performance of the machine learning techniques employed in our study below. Because proposed methodology consisted of two layers for PPD-CC, each layer is discussed separately, and results for both C.V. and holdout set validation are reported.

4.3 | Depressive content classification layer results

Using the feature set, D-CC layer was designed to predict whether a post was depressive or nondepressive in nature. As shown in Table 5, MLPs performed better than SVM and LR for both 10-fold C.V. as well as holdout set. LR and SVM also showed promising results. Consistent performance of machine learning techniques in 10-fold C.V. and the holdout set validation is indicative of the reliability of the feature set. However, it is worth noting that in the domain of natural language processing and for a very large dataset, algorithms like SVM may not be able to perform up to mark. On the other hand, MLPs are generally expected to learn better in the presence of a larger dataset, therefore greatly enhancing their predictive capabilities.

TABLE 5 D-CC performance scores

D-CC layer	Classifier	Accuracy	Precision	Recall
10-fold C.V.	SVM	89.42	89.69	89.40
	MLP	91.63	91.83	91.85
	LR	90.31	90.46	90.30
Holdout	SVM	90.46	90.61	90.50
	MLP	91.70	91.74	91.70
	LR	90.84	90.90	90.87

Abbreviations: C.V., cross-validation; D-CC, depressive content classification; LR, logistic regression; MLP, multilayer perceptron; SVM, support vector machine.

Table 6 shows confusion matrix based on prediction results of MLPs in holdout set validation. Out of 531 depressive posts, 53 were incorrectly classified as nondepressive.

4.4 | PPD content classification layer results

Posts classified as depressive in the first layer were introduced into the second layer. As the name suggests, PPD-CC layer was trained to predict whether a post represented feature values associated with PPD or non-PPD content. Tenfold C.V. results showed a notable difference in the performance of each machine learning techniques used. LR performed the best with average accuracy score of 83.81%, whereas SVM and MLP were able to predict with 80.53% and 76.74% accuracy, respectively (Table 7). Low performance score of MLP can be attributed to the smaller number of PPD posts in training set as compared with non-PPD posts. In case of holdout set validation, MLP outperformed other techniques by achieving accuracy score of 85.51%. Although LR and SVM exhibited decrease in their accuracy score, their results were relatively consistent with their predictive capabilities for 10-fold C.V. Although experiments showed that PPD prediction can be carried out using machine learning with linguistic features, we believe that there is no clear winner among the used algorithms. Although MLP showed better results for holdout validation as compared with 10-fold C.V., due to its inherent ability of learning complex features from an input feature set, it can be expected to yield improved results in the presence of a large training set.

Confusion matrix in Table 8 shows prediction results of MLPs in holdout set validation. From the 512 posts categorized as depressive in D-CC layer, 245 were associated with PPD. MLPs-based model was able to predict 205 correctly, whereas classifier was unable to correctly classify 40 PPD posts, and it also incorrectly categorized 27 non-PPD posts as PPD.

4.5 | Language usage in post titles

We also visualized the frequently used words in the titles of posts for each category to identify the most commonly used terms. We plotted word clouds for each category. Results are shown in Figure 3. Although Figure 3a has a very different set of terms as compared with other two categories, it can be seen that Figure 3b,c has many similar term occurrences. Even for the case of terms occurring in both categories of PPD and

TABLE 6 D-CC layer confusion matrix of MLP for holdout validation

		Predicted	
		depressive	Nondepressive
Actual	Depressive	478	53
	Nondepressive	34	484

Abbreviations: D-CC, depressive content classification; MLP, multilayer perceptron.

TABLE 7 PPD-CC performance scores

PPD-CC layer	Classifier	Accuracy	Precision	Recall
10-fold C.V.	SVM	80.70	82.85	80.00
	MLP	80.36	75.11	80.06
	LR	83.73	84.55	83.36
Holdout	SVM	76.86	79.65	75.60
	MLP	86.91	87.03	86.78
	LR	79.76	80.77	79.01

Abbreviations: C.V., cross-validation; LR, logistic regression; MLP, multilayer perceptron; PPD-CC, postpartum depression content classification; SVM, support vector machine.

TABLE 8 PPD-CC layer confusion matrix of MLP for holdout validation

		Predicted	
		PPD	Non-PPD
Actual	PPD	205	40
	Non-PPD	27	240

Abbreviations: MLP, multilayer perceptron; PPD, postpartum depression; PPD-CC, postpartum depression content classification.

(a) Daily Life (b) Depression (c) PPD

FIGURE 3 Word clouds for post titles. PPD, postpartum depression

(a) Daily Life (b) Depression (c) PPD

FIGURE 4 Word clouds for post contents. PPD, postpartum depression

TABLE 9 Validation of feature set from psychology literature

Feature	Current work	Petrick (1984)	Boyer (1990)	Beck (2001)
1	Family	Emotional support of partner and/or family	Support from family	Social support
			Feeling unloved by partner	Marital relationship
2	Drive	Recent major changes in ones' life	Lack of control of one's life	Self esteem
		Difficulty making changes		
3	Death	Fear of illness		
4	Anger		Angry at your life situation	
5	Home		Financial, housing, or other personal problems	Life stress
				Socio-economic status
6	Negative emotion		Feel it is your fault when bad things happen to you	

depression group, there is a clear difference in frequency of usage. Moreover, we noted that certain frequent words like *baby*, *PPD*, and *birth* only appeared for PPD category, which is quite logical considering the fact that PPD is related to parents and parenthood.

Figure 4 shows word clouds for most frequent words occurring in the posts themselves. Overlap in word usage in the title and the content can be observed by comparing corresponding word clouds, such as Figures 3c and 4c have several terms in common.

4.6 | Discussion

Feature set identified in Section 3.2 covers a wide range of topics, as it predicts three classes, namely, nondepressed, depressed, and PPD. In order to validate the importance of these features in psychology, they are explored in psychology literature. Table 9 shows three different research articles that identified similar features as risk factors for PPD. It can be observed that over the years, researchers have addressed similar factors from different perspectives. Availability of emotional support by family or strength of relationship with the partner has been considered a predictor of PPD in the research. A feature that may relate to this concept on a broad level is "family." However, it should be noted that we do not consider features in our dataset to be exact reflections of the concepts discussed in psychology literature.

Looking at the performance results for each layer, it can be seen that textual features can be used successfully as predictors of mental health to aid in identifying individuals at risk of general depression or a specific form such as PPD. The proposed system recognized 19 features from LIWC-2015 dictionary and an additional feature based on absolutist words' dictionary, as strongest predictors of depressive content in general and PPD in specific. Using machine learning techniques, we can handle large data and yield consistent prediction results as well as further improve the accuracy of our techniques. MLP outperformed other techniques in 10-fold C.V. (91.63% accuracy) and for holdout validation (91.51% accuracy) for D-CC layer, and in PPD-CC layer, it outperformed in case of holdout set (84.17%). Exploring MLPs further could be useful due to the reason that MLPs are able to derive high-level features based on input feature set and have been found to perform better when trained using a large dataset. Moreover, based on our analysis of post titles using word clouds, we can assume that the title of a post can potentially be used to partially predict the category or at least aid in the process.

5 | CONCLUSION AND FUTURE WORK

In this study, we examined the feasibility of employing a combination of text-based features and machine learning techniques to classify depressive and nondepressive posts and then further identify posts representing characteristics of PPD. The evaluation of multiple techniques on posts from social media communities demonstrated the success of proposed methodology. D-CC yielded about 91.7% accuracy, whereas PPD-CC achieved 86.91% accuracy, illustrating strong predictive capabilities of the system.

Because our work explored the possibility of employing textual features, we did not explore the extent to which changes occur in PPD-affected parents. Moreover, we also included PPD posts from parenting-related subreddits because of the low number of posts in PPD-focused subreddits. Multiple factors may be contributing towards the lower number of posts including lack of awareness about these subreddits and the unwillingness of people to share their feelings in a new group as compared with the comfort they might have developed interacting with the subreddits they are already active in. This research does not claim that all the individuals posting in subreddits of interest suffer from PPD; we can only make a weak inference about it while acknowledging the possibility of selective self-presentation. Future intervention studies may be carried out for creation of a large training set containing texts produced in different settings such as social media post, essays, and daily journals. Because we do not address the extent or possibility of causal role of feature set in occurrence of PPD, using time-based content, it can be further studied to identify features as possible vulnerability factors.

ACKNOWLEDGEMENT

This research is supported by the Deanship of Scientific Research, King Faisal University (Project ID 180055).

CONFLICT OF INTERESTS

None.

ORCID

Iram Fatima https://orcid.org/0000-0002-5104-8958

REFERENCES

Al-Mosaiwi, M., & Johnstone, T. (2018). In an absolute state: Elevated use of absolutist words is a marker specific to anxiety, depression, and suicidal ideation. *Clinical Psychological Science*, *6*(4), 529–542. https://doi.org/10.1177/2167702617747074

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders (5th ed.)*. Washington, DC.

Baumel, A. (2015). Online emotional support delivered by trained volunteers: Users' satisfaction and their perception of the service compared to psychotherapy. *Journal of Mental Health*, *24*(5), 313–320.

Beck, C. T. (1998,). The effects of postpartum depression on child development: A meta-analysis. *Archives of Psychiatric Nursing*, *12*(1), 12–20.

Beck, C. T. (2001). Predictors of postpartum depression: An update. *Nursing Research*, *50*(5), 275–285.

Boyer, D. (1990). Prediction of postpartum depression. *NAACOG's Clinical Issues in Perinatal and Women's Health Nursing*, *1*(3), 359–368.

Braverman, J, & Roux, J. (1978). Screening for the patient at risk for postpartum depression. *Obstetrics and Gynecology*, *52*(6), 731–736.

Cox, J. L., Holden, J. M., & Sagovsky, R. (1987). Detection of postnatal depression: Development of the 10-item Edinburgh Postnatal Depression Scale. *The British Journal of Psychiatry*, *150*(6), 782–786.

Davé, S., Petersen, I., Sherr, L., & Nazareth, I. (2010). Incidence of maternal and paternal depression in primary care: A cohort study using a primary care database. *Archives of Pediatrics & Adolescent Medicine*, *164*(11), 1038–1044.

De Choudhury, M., Counts, S., & Horvitz, E. (2013). Major life changes and behavioral markers in social media: Case of childbirth. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*, ACM, San Antonio, TX, USA, pp. 1431–1442.

De Choudhury, M., Counts, S., & Horvitz, E. (2013). Social media as a measurement tool of depression in populations. In *Proceedings of the 5th Annual ACM Web Science Conference*, ACM, Paris, France, pp. 47–56.

De Choudhury, M., Counts, S., Horvitz, E. J., & Ho, A. (2014). Characterizing and predicting postpartum depression from shared facebook data. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing*, Baltimore, Maryland, USA, pp. 626–638.

De Choudhury, M., & De, S. (2014). Mental health discourse on reddit: Self-disclosure, social support, and anonymity. In *Proceedings of the 8th International Conference on Weblogs and Social Media, ICWSM 2014*, Ann Arbor, Michigan, USA, pp. 71–80.

Dimidjian, S., Goodman, S. H., Felder, J. N., Gallop, R., Brown, A. P., & Beck, A. (2015). An open trial of mindfulness-based cognitive therapy for the prevention of perinatal depressive relapse/recurrence. *Archives of women's mental health*, *18*(1), 85–94.

Fatima, I., Mukhtar, H., Ahmad, H. F., & Rajpoot, K. (2017). Analysis of user-generated content from online social communities to characterise and predict depression degree. *Journal of Information Science*, *44*(5), 683–695.

Fletcher, R. (1987). *Practical methods of optimization*. New York, NY, USA: Wiley.

Gardner, M. W, & Dorling, S. R (1998). Artificial neural networks (the multilayer perceptron)—A review of applications in the atmospheric sciences. *Atmospheric Environment*, *32*(14), 2627–2636.

Genesoni, L., & Tallandini, M. A. (2009). Men's psychological transition to fatherhood: An analysis of the literature, 1989–2008. *Birth*, *36*(4), 305–318.

Hanna, R., Rohm, A., & Crittenden, V. L. (2011). We're all connected: The power of the social media ecosystem. *Business Horizons*, *54*(3), 265–273.

Hsu, C.-W., Chang, C.-C., & Lin, C.-J. (2003). A practical guide to support vector classification.

Hudson, D. B., Elek, S. M., & Campbell-Grossman, C. (2000). Depression, self-esteem, loneliness, and social support among adolescent mothers participating in the new parents project. *Adolescence*, *35*(139), 445–454.

Hussain, J., Ali, M., Bilal, H. S. M., Afzal, M., Ahmad, H. F., Banos, O., & Lee, S. (2015). SNS based predictive model for depression. In *International Conference on Smart Homes and Health Telematics*, Geneva, Switzerland, pp. 349–354.

Kalyani, G., Saeed, K, Rehman, C., & Mubbashar, M. (2001). Incidence of depressive illness in pakistani women during postnatal period. *Journal-College of Physicians and Surgeons of Pakistan*, *11*, 246–248.

Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence*, *14*, Montreal, Canada, Montreal, pp. 1137–1145.

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9. *Journal of General Internal Medicine*, *16*(9), 606–613.

Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval*. New York, NY, USA: Cambridge University Press.

Mohr, D. C., Burns, M. N., Schueller, S. M., Clarke, G., & Klinkman, M. (2013). Behavioral intervention technologies: Evidence review and recommendations for future research in mental health. *General Hospital Psychiatry*, *35*(4), 332–338.

Mohr, D. C., Ho, J., Duffecy, J., Baron, K. G., Lehman, K. A., Jin, L., & Reifler, D. (2010). Perceived barriers to psychological treatments and their relationship to depression. *Journal of Clinical Psychology*, *66*(4), 394–409.

Mott, S. L., Schiller, C. E., Richards, J. G., O'Hara, M. W., & Stuart, S. (2011). Depression and anxiety among postpartum and adoptive mothers. *Archives of Women's Mental Health*, *14*(4), 335.

Pearlstein, T., Howard, M., Salisbury, A., & Zlotnick, C. (2009). Postpartum depression. *American Journal of Obstetrics and Gynecology*, *200*(4), 357–364.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., …, Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). Linguistic inquiry and word count: LIWC 2001. *Mahway: Lawrence Erlbaum Associates*, *71*(2001), 2001.

Petrick, J. M. (1984). Postpartum depression identification of high-risk mothers. *Journal of Obstetric, Gynecologic, & Neonatal Nursing*, *13*(1), 37–40.

Reck, C., Stehle, E., Reinig, K., & Mundt, C. (2009). Maternity blues as a predictor of DSM-IV depression and anxiety disorders in the first three months postpartum. *Journal of Affective Disorders*, *113*(1-2), 77–87.

Rude, S., Gortner, E.-M., & Pennebaker, J. (2004). Language use of depressed and depression-vulnerable college students. *Cognition & Emotion*, *18*(8), 1121–1133.

Sarkadi, A., & Bremberg, S. (2005). Socially unbiased parenting support on the internet—A cross-sectional study of users of a large Swedish parenting website, *31*, 43–52.

Schlkopf, B., & Smola, A. J. (2002). *Learning with kernels : support vector machines, regularization, optimization, and beyond*. Cambridge, Mass: MIT Press.

Scholkopf, B., Sung, K.-K., Burges, C. J. C., Girosi, F., Niyogi, P., Poggio, T., & Vapnik, V. (1997). Comparing support vector machines with Gaussian kernels to radial basis function classifiers. *IEEE Transactions on Signal Processing*, *45*(11), 2758–2765.

Shulman, B., Dueck, R., Ryan, D., Breau, G., Sadowski, I., & Misri, S. (2018). Feasibility of a mindfulness-based cognitive therapy group intervention as an adjunctive treatment for postpartum depression and anxiety. *Journal of Affective Disorders*, *235*, 61–67.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, *58*(1), 267–288.

Tikhonov, A.-I. N. (1977). *Solutions of ill posed problems*. Washington, New York: Vh Winston.

Woolhouse, H., McDonald, E., & Brown, S. (2012). Women's experiences of sex and intimacy after childbirth: Making the adjustment to motherhood. *Journal of Psychosomatic Obstetrics & Gynecology*, *33*(4), 185–190. https://doi.org/10.3109/0167482X.2012.720314

World Health Organization (2004). *ICD-10 : international statistical classication of diseases and related health problems / world health organization (10th revision, 2nd ed.)*. Geneva, Switzerland: World Health Organization.

AUTHOR BIOGRAPHIES

Iram Fatima received her PhD in Computer Engineering from Kyung Hee University, South Korea. Currently, she is working as an assistant professor at the College of Computer Sciences and Information Technology, King Faisal University. She has established many research collaborations and secured research grants from Deanship of Scientific Research, King Faisal University, Saudi Arabia. She co-authored more than 30 papers that have been published in international journals and conferences. Her main research interests are in behaviour analysis, social networks, natural language processing, and machine learning. She has also worked on pattern recognition, feature extraction, text analysis, and clinical decision support systems.

Burhan Ud Din Abbasi is a researcher at the Knowledge-Based Systems (KBS) research group at School of Electrical Engineering & Computer Science (SEECS), National University of Science & Technology (NUST), Pakistan. His research interests include natural language processing, machine learning, and deep learning.

Sharifullah Khan received his PhD in Computer Science from the University of Leeds, Leeds, UK, in 2002. He works in School of Electrical Engineering and Computer Science (SEECS), National University of Sciences and Technology (NUST), Islamabad, Pakistan, since 2003. During 2005–2006, he completed his 1-year postdoctoral in the Université Paul Sabatier (UPS), Toulouse, France, which was sponsored by CNRS France. Currently, he is serving as an associate professor at NUST SEECS. He co-authored more than 70 papers that have been published in international journals and conferences. He also leads the Knowledge-Based Systems (KBS) research group at SEECS and secured several funded projects. Dr Khan is conducting research activities in the areas of data science and ontology engineering, and information retrieval.

Majed Al-Saeed is an assistant professor in the Department of Computer Science (CS) at College of Computer Sciences and Information Technology (CCSIT), King Faisal University (KFU), Al-Ahsa, Saudi Arabia. He received his PhD in Computer Science from University of Glasgow, UK, in 2015. He received his MSc in Software Systems Engineering from the University of Melbourne, Australia, in 2018. He received his BSc in Computer and Information Systems from King Faisal University, Saudi Arabia, in 2003. His research interests focus on performance analysis of software systems, in particular measuring, evaluating, and improving performance issues of intelligent systems. His work has been on designing, implementing, and performance evaluation of profilers of parallel domain-specific languages (DSLs) to improving the understanding and performance of high-level distributed parallel functional languages. From 2016 to 2018, he served as the Chairman of CS Department. In 2016, he served as acting Vice Dean of Admissions and Registration in KFU.

Hafiz Farooq Ahmad is an associate professor at the College of Computer Sciences and Information Technology (CCSIT), King Faisal University, Al-Ahsa, Saudi Arabia. He holds a PhD in Distributed Computing from Tokyo Institute of Technology, Tokyo, Japan. He has research interest in semantics systems, machine learning, health informatics, and web application security. He is the pioneer for Semantic Web Application Firewall (SWAF) in cooperation with DTS Inc Japan in 2010. He contributed in agent cites project, a European-funded research and development project for agent systems. He initiated SAGE (Scalable fault tolerant Agent Grooming Environment) project and proposed the concept of decentralized multiagent systems SAGE back in 2002. He has more than 100 international publications including a book on security in sensors. He has been awarded a number of national and international awards such as the Best Researcher Award of the year 2011 by NUST, PSF/COMSTECH Best Researcher of the Year 2005, and Star Laureate Award 2004.

Rafia Mumtaz received her PhD in 2010 and also joined NUST-SEECS in the same year. During her tenure at NUST-SEECS, she has established research collaborations with national and international institutes and government organizations, which include Pakistan Agriculture Research Council, Islamabad, Queen Mary University of London, UK, and University of Leeds, UK. She has also secured research grants from European Space Agency (€50,000), British Council's UK Researcher Links Awards (£4,000), Deanship of Scientific Research, King Faisal University, Saudi Arabia (~70,000 riyal), DPIH, Saudi Arabia (5,000 riyals), NRPU HEC (Rs 3.6 million), and DAAD Germany (approximately €80,000). Her research interests encompass remote sensing, geographical information systems, satellite image processing, and internet of things (IoT).

How to cite this article: Fatima I, Abbasi BUD, Khan S, Al-Saeed M, Ahmad HF, Mumtaz R. Prediction of postpartum depression using machine learning techniques from social media text. *Expert Systems*. 2019;36:e12409. https://doi.org/10.1111/exsy.12409