Early Detection of Workplace Stress Using NLP and Sentiment Analysis on Employee Communication Data

[1]Yashka Johari, [2]Dr.Sakshi
[12]Amity Institute of Information Technology, Amity University Noida, India
[1]yyneha7714@gmail.com, [2]sarora2@gmail.com

***Abstract*—One of the major concerns of the contemporary workplace scenario is the issue of job or work stress, which impacts the wellbeing and success of employees to a great extent. An enormous amount of textual communication data is produced every day owing to the extensive usage of digital communication platforms like social media, email communication, and instant messaging services among others. Information about the emotional state of mind of the employees is conveyed implicitly through digital communication media. Hence, stress prediction and intervention is possible using textual communication data.**

This paper suggests a proposed approach for early prediction of workplace stress using NLP and describes a case study where sentiment analysis and various supervised learning approaches are applied on textual communication data. A total of 8,897 messages are gathered from publicly accessible Twitter and Reddit datasets. As a proxy to represent professional communication data, a set of preprocessing steps including text normalization, lowercase conversion, punctuation removal, stop words removal, tokenization, and lemmatization is conducted to enhance data quality and reduce any language noise in the text corpus. TF-IDF is one of the methods employed to preprocess textual data, making it possible for further processing by the machine learning model. In this research, different types of supervised learning models are examined, namely Support Vector Machines (SVM), Random Forest, Naive Bayes, and Logistic Regression. Among the employed supervised learning algorithms, Random Forest model is suggested to be the best classifier (82.75%), followed by Logistic Regression (82.24%).

***Index Terms*—Natural Language Processing, Sentiment Analysis, Workplace Stress Detection, Machine Learning, TF-IDF, Text Classification.**

I. Introduction

The workplace of today has been significantly affected by the rapid evolution of the usage of digital communication tools such as social media, internal messaging tools, and emails. Massive textual datasets of daily interactions, opinions, emotions, and psychological states of human beings have been generated through these tools. Digital communication tools in today's workplace facilitate easier collaboration and increased work efficiency, but they also increase work intensity, work connectedness, and work-life blurriness. This is because too much stress can have negative impacts on the mental well-being and satisfaction of employees, as well as their productivity. Work-related stress is therefore becoming another major issue in many firms around the globe [1].

Moreover, NLP has emerged as a powerful tool in the automatic analysis of unstructured textual data and extracting significant patterns from it due to the availability of textual data. One of the most popular subdomains of natural language processing (NLP), which has attracted significant interest in recent years, is sentiment analysis. The sentiment analysis of text data is concerned with the attitudes, feelings, and opinions expressed in it. The linguistic characteristics of words and their sentiment polarity can be helpful in

understanding employee communication related to stress using sentiment analysis. The effectiveness of NLP-based methods has already been validated in analyzing mental health using social media data compared to traditional methods of analysis [4].

Because of the capability of these algorithms to detect complex patterns, the use of these algorithms in text classification is becoming more common nowadays. Supervised learning algorithms like Random Forest, Naïve Bayes, Logistic Regression, and SVM have proved successful in building Sentiment Analysis and Stress Detection Applications [5]-[8]. These algorithms can be used effectively to classify the text data into stressful and non-stressful data by using the appropriate features. This paper focuses on introducing a new approach to detect the early stage of stress among workers with the help of textual data derived from workplace communication. According to previous studies conducted on stress detection [10], [11], public data is extracted from well-known social media websites like Twitter and Reddit, which can be utilized to generate workplace data because of its richness in stress, emotion, and opinion. Classification of the data generated with the machine learning algorithm takes place after the preprocessing of data and conversion of textual data into numerical data through TF-IDF.

II. RELATED WORK

The application of these algorithms to text classification is increasingly popular due to their ability to recognize intricate patterns. Supervised learning algorithms including Random Forest, Naive Bayes, Logistic Regression, and Support Vector Machines have been successfully applied to develop sentiment analysis and stress detection systems [5]–[8]. These algorithms can be used effectively to categorize the text data into categories of stress and non-stress with the help of suitable features. This study presents an innovative approach to detect the early signs of stress at work based on text classification using data obtained from workplace communications. Similar to previous research in stress detection [10], [11], publicly available data from popular social media platforms, such as Twitter and Reddit, are used as the source of data for creating workplace data since the data from these sources are highly abundant with stress, emotions, and opinion. Classification of the data with the help of machine learning algorithms is done after pre-processing and transformation of the data into numeric form with TF-IDF.In the domain of Sentiment Analysis, initial concepts were first introduced by Bing Liu [2] for the extraction of opinion and emotions from the text. Such ideas formed the basis of current methods of sentiment analysis. Gerard Salton and Christopher Buckley in their early research in information retrieval [3] introduced many different approaches for term weighting in documents such as TF-IDF. Such approaches continue to play a very important role in feature representation in different machine learning models. Different techniques using machine learning have been used for classification of text. Thorsten Joachims' study [4] illustrates the power of support vector machines in document classification. McCallum and Nigam [6] introduced different probabilistic methods for text classification. The survey conducted by Fabrizio Sebastiani [7] gives an overview of different text classification techniques along with highlighting the potentiality of different machine learning techniques for text classification. Moreover, different ensemble-based techniques such as Random Forest introduced by Leo Breiman [9] have proven to be very effective in classification tasks consisting of multiple decision trees.

Also, advancements have been made in the computation techniques to facilitate the development of machine learning models. One library which has gained popularity because of its efficiency is Python, which has been developed by Fabian Pedregosa et al. in [10]. It has various tools that help in carrying out data processing and modeling. Psychology research work conducted by James W. Pennebaker et al. in [11] reveals that there exists a possibility for associating linguistic components with emotional and psychological factors. In this respect, the works done by Munmun De Choudhury et al. in [12] show that it is possible to detect depression by using social media information. Moreover, other ways of performing sentiment analysis on documents, as mentioned by Qamar Rajput et al. in [13], also offer ways for conducting sentiment analysis on documents.

A. Dataset Description

The chosen database contains 8,897 textual messages which have been collected via Twitter and Reddit for analyzing stress and emotional conditions through textual analysis. Furthermore, the application of sentiment analysis and NLP in recognizing signs of stress in textual analysis is on the rise [2], [3].

The use of machine learning techniques, including Naïve Bayes, Support Vector Machines, and Logistic Regression, in text classification tasks has been shown to be effective because of their simplicity and high precision rates [4]–[7]. Previous studies have also shown that the use of lexical elements and word weighting strategies, such as TF-IDF, have been successful in detecting features associated with stress from text-based data [8], [9]. However, previous studies have also demonstrated that lexical approaches tend to be highly dependent and might be inefficient for dealing with heterogeneous data sets.TF-IDF is a type of word weighting technique where weights are assigned to each individual word according to its frequency within a document and its infrequency in the entire corpus. Words that occur more frequently within a document and less in the entire corpus are assigned higher weights.B. Feature Extraction Using TF-IDF

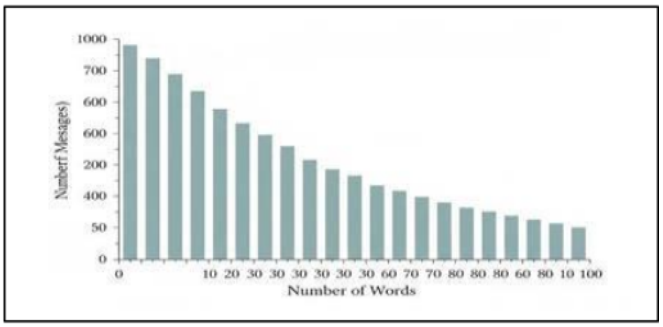

Fig. 1. Distribution of tweets by number of words.

III. PROPOSED METHODOLOGY

This pipeline is an illustration of how various functions are implemented for the purpose of implementing the workplace stress detection framework. Among the functions that have been performed include data gathering, text data pre-processing, feature extraction, and finally classification. The text data gathering has been done using public sites like Reddit and Twitter that have previously been used for purposes of communicating in the workplace due to the rich description of user experiences and emotions [1], [2].

For the current study, stress-related text data and non-stress-related text data are used to carry out the task of supervised learning. Some of the operations carried out on the text data in an attempt to preprocess and improve the quality of the text data include normalization, lowercase conversion, tokenization, lemmatization, punctuation, and removal of stopwords. Lastly, the text data is represented as numeric vectors to prove the significance of text data processing, and the text data is categorized into two groups: stressed and non-stressed text data. This categorization is achieved using different supervised learning techniques such as Logistic Regression, Naive Bayes, Random Forest, and Support Vector Machines.

B. Text Preprocessing

The pre-processing procedures that take place during text pre-processing include lower casing, elimination of punctuations, elimination of stop words, tokenization, and lemmatization.

Lower casing is performed so as to ensure that all the words within the data are consistent in terms of their cases when analyzing or performing machine learning operations on the data. Elimination of punctuations and stop words is performed so as to prevent any unwanted data from appearing in the dataset from the text and eliminate any frequently occurring words within the data set that do not play any important role in the

analysis or machine learning operation performed on the data. Tokenization is performed so as to separate the data into distinct words so as to enable further manipulation of the data.

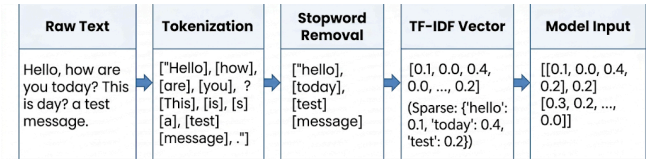

Fig. 2. Text preprocessing steps.

C. Text Preprocessing

Various supervised machine learning classifier models are evaluated and analyzed for establishing the most effective classifier model for stress detection. In this case, the various classifier models that will be used in the analysis include Logistic Regression, Naïve Bayes, Random Forest, and Support Vector Machines (SVM). The Logistic Regression classifier is among the most commonly used classifier models for handling binary classification problems owing to the ease with which it can be understood and applied. On the other hand, the Naïve Bayes classifier model works best in text classification applications. The Random Forest classifier model eliminates the overfitting problems from a classifier model. Support Vector Machines are common classifier models for binary classification problems. This is due to the capability of the classifier model to handle high-dimensional data sets. They can also determine the best decision boundaries for classification between two or more classes in the data set. In this case, the class-weighted Logistic Regression classifier is used for handling class imbalances.

D. System Architecture

As can be seen in Fig. 3, the architecture of the suggested system for detecting workplace stress. As can be seen in Fig., it can be noted that the system consists of several levels, which include data gathering, preprocessing, TF-IDF, classification through a machine learning classifier, and lastly, prediction of the stress. Initially, the system gathers textual data through social media, followed by the preprocessing stage. Afterward, the gathered data are converted to numerical data using TF-IDF. Lastly, the data are classified via a machine learning classifier and predict the stress level in the gathered textual data.

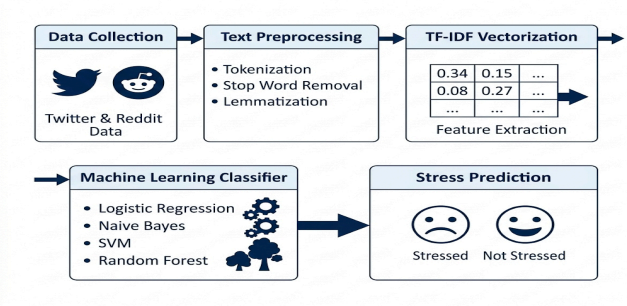

Fig. 3. Overall system architecture for workplace stress detection using machine learning.

IV. EXPERIMENTAL SETUP AND EVALUATION

A. Train–Test Split

As illustrated in Fig. 3, the system architecture of the proposed framework for workplace stress detection. As can be seen from the above figure, it can be noted that the system follows several stages, such as data collection, data preprocessing, TF-IDF, classification with the help of the machine learning classifier, and, finally, stress prediction. The system begins with data collection from social media platforms and continues with data preprocessing. Following the data preprocessing stage, it continues with the

conversion of the data into numerical form by using TF-IDF, and, finally, it proceeds with classification with the help of machine learning classifier and stress prediction. Initially, the system collects textual data from social media platforms, followed by the data preprocessing stage.

B. Evaluation Metrics

For the evaluation of the performance of the suggested model for detecting stress, several measures for the evaluation of classification performance are employed, including Accuracy, Precision, Recall, and F1-Score. The aforementioned measures offer a holistic assessment of the classifier's performance.

Accuracy is defined as:

Precision is defined as the proportion of correctly predicted positive cases out of all predicted positive cases:

Recall is defined as the proportion of correctly predicted positive cases out of all positive cases in the dataset:

The F1-Score is the harmonic mean of precision and recall. The importance of using the F1-score when evaluating stress detection models lies in the fact that the F1-Score provides an objective view of the performance of a classifier by considering both precision and recall. $$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

V. RESULTS AND DISCUSSION

A. Model Performance Comparison

The performance of the various classifiers is evaluated using the TF-IDF feature extraction method, and the following results were obtained:

The highest accuracy of 82.75% was recorded for the Random Forest Classifier, while the accuracy of the Logistic Regression Classifier was 82.24%.

Model	Accuracy (%)	Precision	Recall	F1-Score
Random Forest	82.75	0.83	0.82	0.82
Logistic Regression	82.24	0.82	0.82	0.82
SVM	81.01	0.81	0.81	0.81
Naïve Bayes	79.04	0.79	0.79	0.79

TABLE I : MODEL PERFORMANCE COMPARISON

For the determination of workplace stress using textual communication data, four different supervised machine learning algorithms such as Logistic Regression, Naïve Bayes, Random Forest, and Support Vector Machine (SVM) were applied and evaluated. In all these models, the TF-IDF features were used as a fixed dimension of 1,000. For proper evaluation of the performance of the models, the dataset was divided into 80:20 for the training and testing phase.

Accuracy, precision, recall, and F1 score metrics were utilized to evaluate the performance of these various machine learning algorithms. Out of the mentioned four algorithms, the performance of the Random Forest

model was higher in comparison to other models in terms of accuracy (82.75%). The reason behind this is that the model consists of multiple decision trees which enhance the overall performance of the model.

The performance of the logistic regression model is comparatively lower than the Random Forest but still it can be considered.

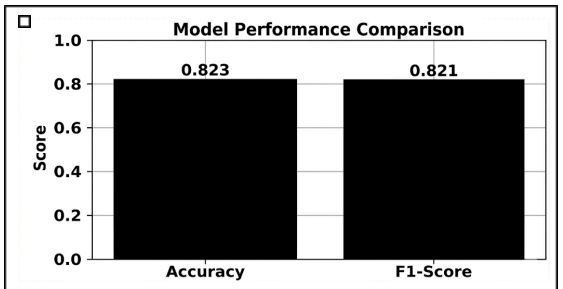

Fig. 4. Logistic regression performance.

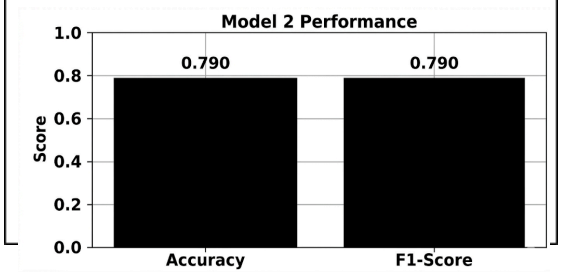

Fig. 5. Naïve Bayes performance.

The structure of the proposed workplace stress detection system is illustrated in Fig. 3. After the collection of text data from social media websites such as Reddit and Twitter, preprocessing methods are used to preprocess this data. The text is then converted to various feature vectors through TF-IDF vectorization. Finally, different classification techniques are used to check whether or not this communication contains any stress.

As shown in Fig. 4, the performance of the Logistic Regression model is assessed using accuracy and F1-score metrics. Based on Fig. 4 above, it is clear that the performance of the Logistic Regression is quite high at 0.823 for accuracy and 0.821 for F1-Score. This means that Logistic Regression model can identify stressed patterns in communication.

It can be clearly seen that Fig. 5 above shows the performance achieved using the Naïve Bayes classification algorithm. As seen in the graph above, the algorithm achieves an accuracy and F1-score of about 0.790. Although slightly less compared to the one achieved by Logistic Regression, the classifier still has a good performance level in classifying stressed texts.As shown in Figure 5 below, it it is evident that the model is capable of achieving a good accuracy score and an F1-score that is just under 0.790. It may be slightly inferior compared to the one obtained from using logistic regression, but it is evident that it performs moderately well in differentiating stressed from non-stressed texts.

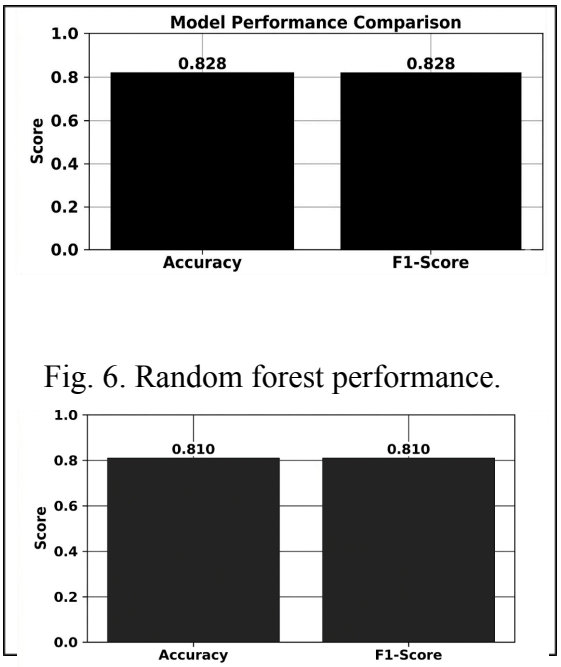

Fig. 6. Random forest performance.

Fig. 7. Linear SVC performance.

B. Linguistic Indicators of Stress

Based on the analysis above, it is clear that messages relating to stressful situations are always accompanied by words like 'tired,' 'stressed,' 'anxious,' and 'depression.' These words can be found in textual messages in which a person is feeling negative emotions and uncomfortable states. On the other hand, messages from persons in a situation that is not stressed are usually accompanied by words such as 'happy,' 'excited,' and 'calm,' which are words that reflect positivity.

In this case, the presence of these words in the messages shows how textual messages can provide insightful information about the mental condition of a person. With the help of words such as these and stressed patterns in messages, stressed and non-stressed states of persons can be differentiated easily using the suggested machine learning model.

C. Limitations and Ethical Considerations

For example, in this research, social media is being used as a representation of the means of communication at work. Nonetheless, social media does not represent the communication taking place in a typical work environment. Even though there are numerous amounts of textual data available for use on social media platforms like Twitter and Reddit, they do not necessarily serve as a reflection of communication at work. Furthermore, certain ethical implications will have to be considered during the application process in a real-life scenario. For example, issues of privacy and communication bias regarding stress arise.

VI. CONCLUSION AND FUTURE WORK

In the present paper, a comprehensive framework has been proposed for the early detection of workplace stress using natural language processing techniques. The results clearly indicate the effectiveness of the proposed framework, in which the classifier has been found to produce the highest accuracy of 82.75% using the Random Forest classifier and TF-IDF features for text representation. This clearly indicates the effectiveness of traditional machine learning approaches, along with the use of

well-engineered textual features, for the detection of the early symptoms of stress among employees. In addition, the results also clearly indicate the effectiveness of the Logistic Regression classifier for the proposed approach, which is also offering good results in terms of achieving the right balance between accuracy and interpretability, especially for high-dimensional text data. It should be noted that the results are obtained using publicly available social media data, which may not be representative of real communication data used in the workplace.

For future work, this proposed framework can also be enhanced further to include the efficiency of recently introduced deep learning methods, such as Long Short-Term Memory (LSTM) and Bidirectional Encoder Representations from Transformers (BERT), in recognizing complex patterns in textual communication data. Furthermore, this proposed framework can also be enhanced to recognize multi-language work-related stresses and be tested using organizational communication data. Further enhancement of this proposed framework can also improve the accuracy of recognizing work-related stresses, which can eventually lead to intelligent systems to monitor employee wellness and develop a healthy work environment.

References

[1] A. Saleem, et al., "Workplace stress and its impact on employee performance," Int. J. Manag. Stud., vol. 2, no. 3, pp. 45–52, 2015.

[2] B. Liu, Sentiment Analysis: Mining Opinions, Sentiments, and Emotions, Morgan & Claypool Publishers, 2015.

[3] G. Salton and C. Buckley, "Term-weighting approaches in automatic text retrieval," Information Processing & Management, vol. 24, no. 5, pp. 513–523, 1988.

[4] T. Joachims, "Text categorization with support vector machines," in Proc. European Conf. on Machine Learning, 1998, pp. 137–142.

[5] Y. Goldberg, Neural Network Methods for Natural Language Processing, Morgan & Claypool, 2017.

[6] A. McCallum and K. Nigam, "A comparison of event models for Naïve Bayes text classification," in AAAI Workshop on Learning forText Categorization, 1998.

[7] F. Sebastiani, "Machine learning in automated text categorization,"

[8] B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up? Sentiment classification using machine learning techniques," in Proc. EMNLP, 2002, pp. 79–86.

[9] L. Breiman, "Random forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.

[10] F. Pedregosa et al., "Scikit-learn: Machine learning in Python," J. Mach. Learn. Res., vol. 12, pp. 2825–2830, 2011.

[11] J. Pennebaker, M. R. Mehl, and K. G. Niederhoffer, "Psychological aspects of natural language use," Annual Review of Psychology, vol. 54, pp. 547–577,2003.

[12] M. De Choudhury, S. Counts, and E. Horvitz, "Predicting depression via social media," in Proc. ICWSM, 2013, pp. 128–137.

[13] Q. Rajput, S. Haider, and S. Ghani, "Lexicon-based sentiment analysis of text documents," J. Appl. Comput., vol. 12, no. 3, pp. 45–54, 2016.