Mental Health: Exploring Neural Networks To Study Stress In The Tech Industry

Sindhu Subramanya
(imanya@stanford.edu)

Abstract

"Rapid economic change is one of the most significant aspects of our times. Behind the curtain of world economy globalization, the occupational environment is delivering increasing stress, such as job insecurity, increasing workload, and the burden of lifelong learning. Larger proportion of employees became involved in service- and knowledge-based industries requiring heavy technological preparedness and mental stress. If the pace of change exceeds the capacity of the workers to cope, negative stress reactions can occur. Negative stress reactions are not limited to the individual worker, but also may affect the worker's family, and the community in which they live." Ref [4]-[10].

One of the main challenges is timely identification of the condition, given the stigma around mental health conditions. In this work, we strive to bridge that gap by exploring unique ways to identify stress and quantify it using neural networks. Our final aim is to find a combination of identifiers which can be reliably used alongside artificial intelligence and neural networks to predict stress levels and burnout preemptively.

INTRODUCTION

"Workplace can function as a secure base for promoting mental health since many workers spend more awake time at work than at home. But the stigma attached to having a psychiatric disorder is such that employees may be reluctant to seek treatment. It is desirable to focus on how to get rid of any potential risk rather than who is most responsible for it. Because there are no clear biological diagnoses, psychological disorders are difficult to diagnose. Currently, it is purely on the basis of emotional states and clinical observations of the behaviors that the individual engages in." Ref [4]-[10].

A detailed study of sections *1 A)-D)* demonstrates gaps and room for exploration in the following areas:

1) A method to derive mapping from the symptoms defined in psychology handbooks to day-day functioning in the high technology workplace environment.

2) An apparatus to ease the diagnostics and make it seamless such that a person's indicative symptoms are diagnosed even before the 4 D's Deviance, Dysfunction, Distress, Danger are encountered.

This need can potentially be addressed by innovative methods of artificial intelligence and deep learning neural networks in a non intrusive manner.

In this paper, an apparatus based on RNN LSTM is used to measure mood changes as an indicator. A limited mapping of various factors in day-day life of a tech worker is listed out. Also, a method to derive a "scorecard" for stress is being experimented.

I. DEFINING PSYCHOLOGICAL DISORDER

The following four sections are derived from ref [4]-[10] and provide the details of what the occupational stress model looks like, methods to diagnostic methods for mood disorders and classification of stress symptoms.

A. Bio-Psycho-Social Model

"The bio- psycho-social model of disorder proposes that disorders are caused by biological, psychological, and social-cultural factors.", as explained in Ref [4].

B. Model of Occupational Stress and Mood Disorder

In Ref [7], the relationship between occupational environment and mood disorders is explained. It details how vicious cycles of "absenteeism" and "presenteeism" are created.

C. Diagnostic methods for conditions related to mood disorder

Further, in Ref [7], the diagnostic methods for conditions related to mood disorders is detailed.

D. Symptoms to identify stress

Detailed classification and symptoms to identify stress are presented in Ref [5]. They are used as a guiding factor to pick and choose identifiers for the input data to our neural network.

II. RELATED WORK

Although there are several studies being conducted in the areas of employing artificial intelligence for mental health, high technology workers are a microcosm of the population with certain unique characteristics and it is well worth studying this sample set separately.

MIT researchers [11] detail a 'context-free' neural-network model that can be unleashed on raw text and audio data from patient interviews to discover speech patterns indicative of depression. In [1],[2], Stanford students discuss neural network systems that can detect 8 major dimensions and basic emotions from speech. It was created using the SEMAINE dataset. Reference [3] talks about an experiment to use smartphones to track the progress of patients with bipolar disorder in Austria. [12] details an interesting implementation of a laugh detector.

III. DATASET AND FEATURES

(i) Dataset: X_train,Y_train (127 sentences for training)

(ii) Y_train is integer label between 0 and 4 corresponding to an emotion for each sentence.

(iii) X_train X_test, Y_test (56 sentences for testing)
30 minute audio recording "A" of tech worker speaking at work.

(iv) Audio->Text conversion "A"->"P". "P" is paragraph of sentences.

(v) Split "P" into sentences "S". This is input to Neural Network.

(vi) Pre-trained word embeddings for dictionary of 400,001 words is used.

(vii) First the training sentences are converted into lists of indices, zero-padded so that their length is the length of the longest sentence.

(viii) This is to ensure all the mini-batches are of the same length.

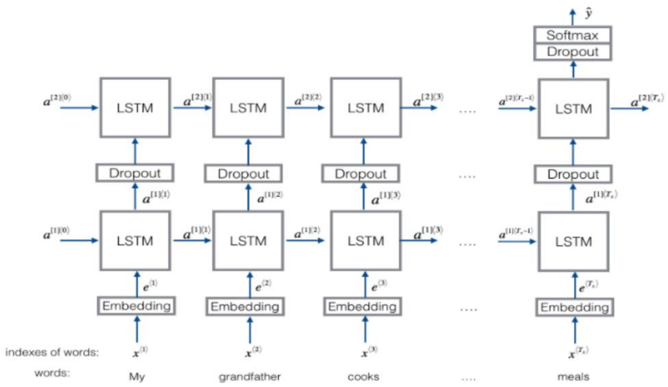
(ix) The pre-trained Embedding layer using the embedding matrix is defined and made non-trainable.

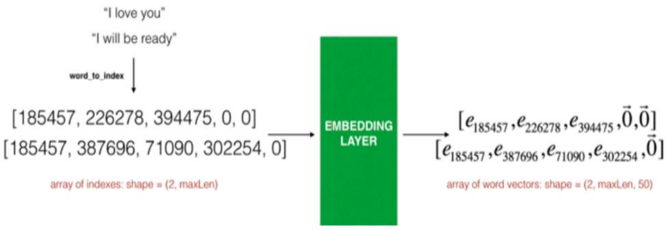
(x) Output of Embedding layer is connected to LSTM network which takes word sequences as input.

(xi) LSTM model takes word ordering into account and predicts the most appropriate emotion conveyed by the sentence.

(xii) Keras is used to compile the model, define the categorical cross entropy loss, Adam optimizer and accuracy metrics.

IV. METHODS






```
Layer (type)                 Output Shape              Param #
=================================================================
input_1 (InputLayer)         (None, 10)                0
_________________________________________________________________
embedding_2 (Embedding)      (None, 10, 50)            20000050
_________________________________________________________________
lstm_2 (LSTM)                (None, 128)               91648
_________________________________________________________________
dropout_2 (Dropout)          (None, 128)               0
_________________________________________________________________
dense_1 (Dense)              (None, 5)                 645
_________________________________________________________________
activation_1 (Activation)    (None, 5)                 0
=================================================================
Total params: 20,092,343
Trainable params: 92,293
Non-trainable params: 20,000,050
_________________________________________________________________
```

Figure 1,2,3 - Model details as seen in Keras

V. RESULTS

(i) Model is trained to take as input an array of shape (m, max_len) and output probability vectors of shape (m, number of classes).

(ii) So X_train (array of sentences as strings) is converted to X_train_indices (array of sentences as list of word indices), and Y_train (labels as indices) to Y_train_oh (labels as one-hot vectors).

(iii) Epochs 50, Batch size = 32.

(iv) Test Accuracy using (X_test, Y_test) of 56 sentences: > 90%.

(v) Now the model is used to interpret "S".

(vi) Output of predicted emotions and changes is plotted accurately.

(vii) The labels used in this experiment are as follows:
0 - Heart, 1 - Sport/Other, 2 - Happy, 3 - Unhappy, 4 - Hungry

VI. CONCLUSION

(i) As expected, the sentiment changes in a monologue conversation 30 minute conversation was captured using this simple RNN LSTM and plotted. So it was proved with a simple example that an RNN can be used to track mood changes.

(ii) By extending the training set and pre-trained word matrix to include a larger set, it is possible to capture most of the speech words. Also, as suggested by the project mentor, transfer learning can be experimented with.

(iii) By plugging in always on speech-text converters, it is possible to get a steady stream of text.

VII. FUTURE WORK

The challenge for most companies is that stress relief programs are purely voluntary and left to the individual's choice to avail them. It is very difficult for an individual to predict their tolerance level and seek timely assistance. This is where AI can help.

If we can think of unique ways to bridge this gap using AI and Deep Learning, it would make a significant and impactful difference to everyone. The derived analysis can be used in various ethical ways to improve the quality of life for everyone.

Based on the author's analysis, listed in below table (Table 1) are some markers that she feels can be used as good indicators to monitor the stress level of tech workers. In this project, "1.Mood and emotional disposition" has been used as an example coupled with an RNN apparatus. Similarly, other markers can be mapped to complementing neural network models to quantify them.

After studying the other indicators, we can represent the findings per person in regularized format.

Example: A "card" with *100 X 100 X Time* parameters similar to image pixels representing various markers per person.

Using these "cards" as input to deep learning, we should be able to identify novel patterns and:

(i) Represent the stress level of a tech worker on a scale (such as 1 through 10)

(ii) Predict the tipping point based on duration for which a tech worker has been exposed to various degrees of stress.

(iii) Approximate the life expectancy or health risks associated with each stress level.

(iv) Predict when to take prevention measures to help avoid burnout.

(v) Predict the recovery period from a burnout condition given certain measures were taken.

Unique indicators for a Person	***Unique indicators in an office area/team***
• Mood and emotional disposition • Confidence and general personality disposition • Posture, gait, weight and physical disposition • Time spent in sunlight per day • Time spent in fresh air per day • Travel time, frequency of travel and jet lag • Holidays availed for relaxation • Speech/Lack of speech • Eyes strain and screen time • Keyboard typing style • Work hours onsite and remote • Tobacco/Alcohol/Substance consumption • Caffeine consumption • Water consumption • Food habits • Exercise • Driving patterns • Sleep patterns • Alternately abled • History of life events • Recent life events	• Indoor lighting condition • Indoor air quality • Ergonomic conditions • Measure laughter in teams. How often teams laugh together can be measured using deep learning (reference link 11) • Team dynamics and energy • Work environment and quiet spaces to think, work, rest. • Social/Political/Economic climate

VIII. CODE

Git Hub Link: https://github.com/simanyas/CS230

IX. ACKNOWLEDGEMENT

The author would like to thank the project mentor Tugce Tasci, Stanford University and Professors Andrew Ng, Kian Katanforoosh, Stanford University.

Also special thanks to Sanjana Murthy, Johns Hopkins University for her insights and consultation on psychology.

X. REFERENCES

1. A.Balakrishnan and A.Rege, “Reading Emotions from Speech using Deep Neural Networks” , Stanford University, CA, June 2017.
2. R.Pryzant, “Context-Aware Response Generation for Mental Health Counseling” , Stanford University, June 2017.
3. Osmani, Gruenerbl, Bahle, Haring, Lukowicz, Mayora, “Smartphones in Mental Health,”, Sept 2015
4. https://opentextbc.ca/introductiontopsychology
5. https://isma.org.uk/about-stress/how-to-identify-stress
6. https://www.medicalnewstoday.com/articles/314093.php
7. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2559945/
8. https://www.ncbi.nlm.nih.gov/books/NBK64063/
9. https://www.health.harvard.edu/newsletter_article/mental-health-problems-in-the-workplace
10. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5483785/
11. http://news.mit.edu/2018/neural-network-model-detect-depression-conversations-0830
12. https://labs.ideo.com/2018/06/15/how-to-build-your-own-laugh-detector/