

Artificial Neural Network Classification Models for Stress in Reading

Nandita Sharma and Tom Gedeon

Research School of Computer Science
Australian National University, Canberra, Australia
{Nandita.Sharma, Tom.Gedeon}@anu.edu.au

Abstract. Stress is a major problem facing our world today and it is important to develop an understanding of how an average person responds to stress in a typical activity like reading. The aim for this paper is to determine whether an artificial neural network (ANN) using measures from stress response signals can be developed to recognize stress in reading text with stressful content. This paper proposes and tests a variety of ANNs that can be used to classify stress in reading using a novel set of stress response signals. It also proposes methods for ANNs to deal with hundreds of features derived from the response signals using a genetic algorithm (GA) based approach. Results show that ANNs using features optimized by GAs helped to select features for stress classification, dealt with corrupted signals and provided better classifications. ANNs using GAs were generated to exploit the time-varying nature of the signals and it was found to be the best method to classify stress compared to all the other ANNs.

Keywords: stress classification, artificial neural networks, genetic algorithms, physiological signals, physical signals, reading.

1 Introduction

Stress was coined by Hans Selye and he defined it as “the non-specific response of the body to any demand for change” [1]. It is the body’s reaction or response to the imbalance caused between demands and resources available to a person. Stress is seen as a natural alarm, resistance and exhaustion [2] system for the body to prepare for a fight or flight response to protect the body from threats and changes. When experienced for longer periods of time and not controlled, stress has been widely recognized as a major growing concern in our age adversely impacting society due to its potential to cause chronic illnesses (e.g. cardiovascular diseases, diabetes and some forms of cancer) and high economic costs in societies (especially in developed countries [3, 4]). Benefits of stress research range from improving personal operations, through increasing work productivity to benefitting society - motivating interest, making it a socially beneficial area of research and posing technical challenges in Computer Science. Numerous computational methods have been used to objectively define and classify stress to differentiate conditions causing stress from other conditions. The methods used simplistic models formed from techniques like

Bayesian networks [5], decision trees [6] and support vector machines [7]. These models have been built from a relatively smaller set of stress features than the sets used in the models in this paper.

The human body's response signals obtained from non-invasive methods that reflect reactions of individuals and their bodies to stressful situations have been used to interpret stress levels. These measures have provided a basis for defining stress objectively. Stress response signals used in this paper fall into two categories - physiological and physical signals. The physiological signals are the galvanic skin response (GSR), electrocardiogram (ECG) and blood pressure (BP). Unlike physiological signals, we define physical signals as a time-varying characteristic where changes can be seen by humans without the need for equipment and tools that need to be attached to individuals to detect general fluctuations. However, sophisticated equipment and sensors using vision technologies are still needed to obtain physical signals at sampling rates sufficient for data analysis and modeling. The physical signals used in this paper are eye gaze positions and pupil dilation. GSR, ECG, BP, eye gaze tracking and pupil dilation have been used to detect stress in literature [5, 8, 9] but this combination is novel to stress research. In this paper, we refer to the physiological and physical signals as primary signals for stress.

Artificial neural networks (ANNs) have been successfully used for emotion-based stress classification for playing video games [10] and preliminary work in stress classification for reading [11]. This paper uses ANN models to recognise *stressed* and *non-stressed* reading based on features derived from primary measures for stress.

Hundreds of stress features can be derived from primary signals for stress to classify stress classes for the different types of reading. If an ANN used all these features as inputs then it would result in quite a large network. This can lead to the issue where the ratio for number of samples to the number of connections and weights in an ANN could be relatively small, which could affect classifications. Furthermore, in order to achieve a good classification model, it should be robust to input tuples or features that suffer from corruption. A genetic algorithm (GA) could help solve these problems by selecting subsets of features for optimizing ANN stress classifications. A GA is based on the concept of natural evolution and has been commonly used to solve optimization problems [12]. It evolves a population of candidate solutions using crossover, mutation and selection methods in search for a population of a better quality. The quality for each individual or chromosome in the population is defined by some fitness function. GAs have been successfully used to select features derived from physiological signals to classify emotions [13, 14].

ANNs have been used in microarray studies for gene expression levels for classification. Like the stress feature space used in this paper, they also faced the issue of dealing with a relatively large feature space [15], tens of magnitude larger than our feature space and with fewer samples. A GA was used to optimize the inputs for an ANN by maintaining features that produced better classifications [15]. The size of a chromosome in the GA, which mirrored the number of features and the number of inputs to the ANN, was set based on what was previously proposed in the literature regarding a similar type of problem [16]. The chromosome size was fixed. For our stress features, we need to discover the number of features that is sufficient for stress

classification. This paper investigates an ANN with all features as inputs (which is referred to as ANN-AllInputs), another ANN with inputs selected by a GA to use features that produced better classifications (ANN-GAInputs) and an ANN with relatively fewer feature inputs and inputs defined to exploit the time-varying characteristic for features (ANN-3Seg-GAInputs). The performance of the ANNs will not only provide how well stress can be classified from the features derived from the primary signals for stress but also provide an insight to irrelevant features that exist in the feature space.

This paper presents the method for data collection from the reading experiment. Various primary signals for stress were recorded during the experiment, which were GSR, ECG, BP, eye gaze tracking and pupil dilation signals. The paper then describes the features obtained from the signals that were used by the different classification models to learn patterns to classify the different types of reading. As mentioned above, the three different types of classification models were ANN-AllInputs, ANN-GAInputs and ANN-3Seg-GAInputs. Each model made use of different facets of stress feature signals from determining features that produced better stress classifications to exploiting time-varying characteristics for signals. Further, the paper provides results, analyses the results and presents a discussion for the results. It concludes by summarizing the work and proposes work that can be done in future to extend the research.

2 Data Collection from Reading Experiment

Thirty-five Undergraduate Computer Science students, comprising 25 males and 10 females, over the age of 18 years old were recruited as experiment participants (after obtaining Ethics Approval from the Australian National University Ethics Committee). Each participant had to understand the requirements of the experiment from written experiment instructions with the guidance of the experiment instructor before they filled in the experiment consent form. Afterwards, physiological stress sensors were attached to the participant and physical stress sensors were calibrated. The instructor notified the participant to start reading, which triggered a sequence of text paragraphs. After finishing the reading, participants had to do an assessment. The experiment process was an extension to the experiment process done in [11] in that the experiment was extended to include a wide range of stress sensors whereas the earlier experiment only had the GSR sensor. An outline of the process of the experiment for an experiment participant is shown in Fig. 1.



Fig. 1. Process followed by participants during the reading experiment

Each participant had physiological and physical measurements taken over the 12 minutes reading time period. During the reading period, a participant read *stressed* and *non-stressed* types of text. Stressed text had stressful content in the direction towards distress, whereas the non-stressed text had content that created an illusion of meditation or soothing environments. Each type of text had the same number of paragraphs and each paragraph was displayed on a computer monitor for participants to read. For consistency, each paragraph was displayed on a 1050 x 1680 pixel Dell monitor, displayed for 60 seconds and positioned at the same location of the computer screen for each participant. Each line of the paragraph had 70 characters including spaces.

Feature values were derived from physiological and physical signals. Biopac ECG100C, Biopac GSR100C and Finapres Finger Cuff systems were used to take ECG, GSR and blood pressure recordings at a sampling rate of 1 kHz. Eye gaze and pupil dilation signals were obtained using Seeing Machines FaceLAB system with a pair of infrared cameras at 60 Hz. Other signals were derived from primary signals such as, heart rate variability, which was calculated from consecutive ECG peaks and another popular signal used for stress detection [17, 18]. Statistics (e.g. mean and standard deviation) were calculated for the signal measurements for each 5 second interval during the stressed and non-stressed reading. Measures such as the number of peaks for periodic signals, the distance an eye covered, the number of forward and backward tracking fixations, and the proportion of the time the eye fixated on different regions of the computer screen over 5 second intervals were also obtained. The regions of the computer screen are shown in Fig. 2. The statistic and measure values formed the stress feature set. In total, there were 215 features.

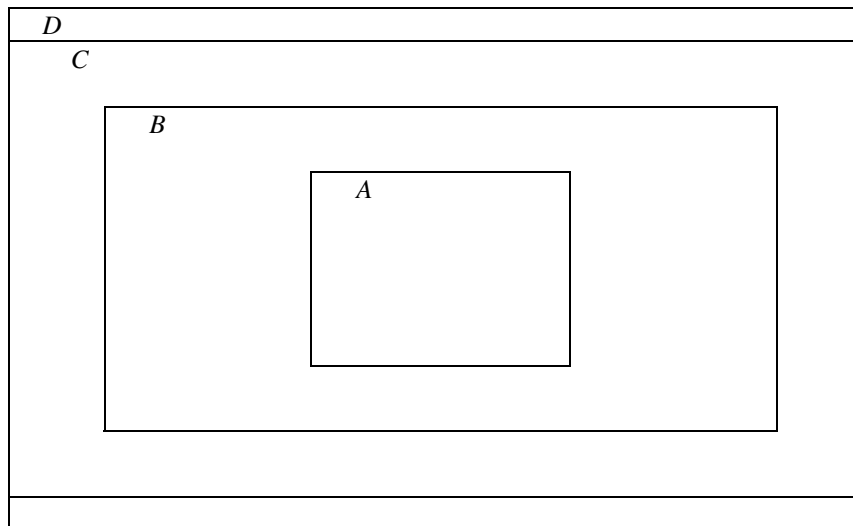


Fig. 2. The bounding rectangles defined the different regions of the computer screen to determine a subset of eye gaze features. The bounding rectangles show regions A, B, C and D. Region A contained the text area where text was displayed to a participant. Region B is the region without A and regions C and D are defined similarly. Region D had the application menu-bar and the tool-bar. Note that the diagram is not drawn to scale.

3 Artificial Neural Network Stress Classifiers

ANNs, inspired by biological neural networks, have characteristics for learning patterns to classify input tuples into classes. It is made up of interconnected processors, known as *artificial neurons*, which are connected by weighted links that pass signals between neurons. In this paper, feed-forward ANNs trained using backpropagation were used. Three topologies were used to classify stress in reading. For two ANNs, a GA was used to select stress features, which formed inputs to the ANNs. As a consequence, the three ANNs differed only on the number of inputs used. The ANNs were:

ANN-AllInputs: A feed-forward ANN was generated for each participant. The ANN had 215 inputs, which mirrored the number of features derived from the primary signals for stress.

ANN-GAInputs: This ANN was similar to ANN-AllInputs but the number of inputs was based on the features selected by a GA. The GA selected features to improve the classification rate for stressed and non-stressed states.

ANN-3Seg-GAInputs: An ANN was generated for each participant with at most 10 features with 3 sequential time segments for each feature as inputs to the ANN. Like ANN-GAInputs, the GA selected features to improve the classification rate for stressed and non-stressed states.

GAs were not only used to improve the classification by selecting relevant features but also to leave out corrupted features. For instance, ECG sensors malfunctioned when acquiring ECG signals for participant with index 19, and this corrupted corresponding features. This information was recorded during data acquisition and can also be seen from observing graphs of raw signals. With a GA, there is a better chance that the ANN classifiers will not use corrupted features for developing relationships from features for stress classification.

A reason for selecting fewer features than the total number of features for ANN-3Seg-GAInputs was to determine whether fewer features can be used to successfully represent the feature space with a smaller network. In addition, if all the features were used as inputs then, in total, the number of inputs would have been 3-fold greater than ANN-AllInputs. Results from ANN-GAInputs show that all 215 features were not needed for an ANN classification. Using ANN-GAInputs, it was found that a smaller subset of features can produce a classification with a better accuracy. This provided another motivation to use fewer features as inputs in ANN-3Seg-GAInputs.

The data sets for each participant was divided up into 3 subsets – training, validation and test sets – where 50% of the data samples were used for training the ANN and the rest of the data set was divided up equally for validating and testing the ANN.

MATLAB was used to implement and test the ANNs. The MATLAB `adapt` function was used for training the ANN on an incremental basis. Each network was trained for 1000 epochs using the Levenberg-Marquardt algorithm. The network had 7 hidden neurons and one neuron in the output layer. Future work could investigate

optimizing the topology of the ANN for stress classification on the reading data set. The accuracy, sensitivity, specificity and the F-Score were calculated to determine the quality of the classification.

GAs were used to select features in ANN-GAInputs and ANN-3Seg-GAInputs based on quality measures for stress classification from ANNs. A GA is a global search technique and has been shown to be useful for optimization problems. Given a population of subsets of features, the GAs evolved the feature sets by applying crossover, mutation and selecting feature sets during each iteration of the search to determine sets of features that produced better quality ANN classifications. The initial population for the GAs was set up to have all the features. The number of features in the chromosomes varied but the chromosome length was fixed. The length of a chromosome was equal to the number of features in the feature space. A chromosome was a binary string where the index for a bit represented a feature and the bit value indicated whether the feature was used in the ANN. For ANN-GAInputs, the initial population had features varying in numbers equally distributed from 1 to 215 features. On the other hand for ANN-3Seg-GAInputs, each chromosome in the initial population had 5 features with feature values for 3 consecutive time segments. The other parameters for the GA were set as provided in Table 1.

Table 1. Settings for GA parameters for ANN-GAInputs and ANN-3Seg-GAInputs

GA Parameter	Value/Setting
population size	100
number of generations	2000
crossover rate	0.8
mutation rate	0.01
crossover type	MATLAB's Scattered Crossover
mutation type	MATLAB's Uniform Mutation
selection type	MATLAB's Stochastic Uniform Selection

4 Results and Discussion

Stressed and non-stressed states for participants were classified using ANN-AllInputs, ANN-GAInputs and ANN-3Seg-GAInputs. The accuracy values for the classification are provided in Fig. 3 for each experiment participant. The best accuracy values were obtained from ANN-3Seg-GAInputs. The accuracy values were significantly different for the three ANNs based on the Student's T-test ($p < 0.01$).

ANN-AllInputs performed the worst compared to the other two ANNs. This shows that the ANN is susceptible to features that were irrelevant and redundant for stress classification. The performance for ANN-GAInputs, which was better than ANN-AllInputs, confirmed this. It used features that were optimized for stress classification. The ANN that exploited the time-varying nature associated with stress features gave the best stress classification. Results were significantly better for ANN-3Seg-GAInputs. This showed that feature values for stress has information in the patterns over time – a significant observation derived directly from this ANN result.

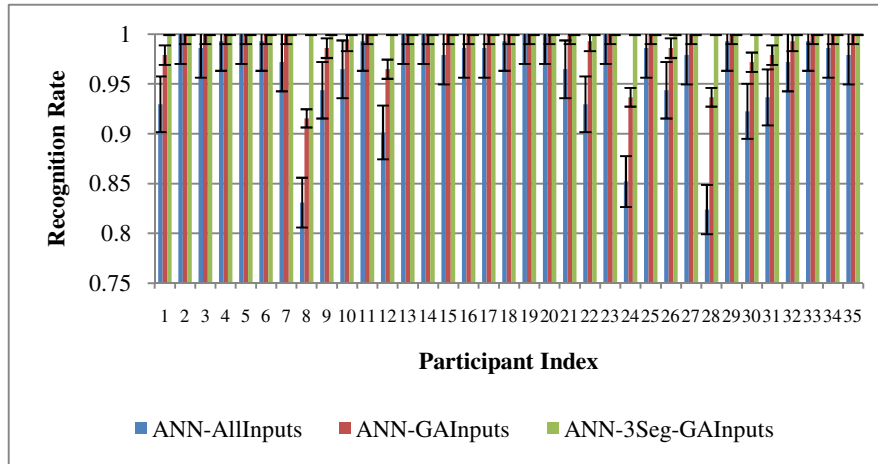


Fig. 3. Recognition rates obtained from cross-validation on the different ANNs

The classification results from the ANNs mirrored the participant reported responses for stress for the different types of text. Participants classified stressed and non-stressed text where classifications were significantly different based on the Chi-test ($p < 0.01$).

5 Conclusion and Future Work

Stress in reading has been successfully classified using an artificial neural network that was built on features derived from an individual's stress response signals. Features can become voluminous and this may increase chances for an ANN to use irrelevant and redundant input features. As a consequence, a genetic algorithm was developed for selecting relevant features for the ANN classification. The GA and ANN hybrid provided better quality classifications. Moreover, by exploiting the time-varying nature of features and incorporating it in the hybrid, the quality of the stress classifications improved and formed a more robust stress classification model. Investigations in the future could examine how different ANN models and topologies may influence stress classification, and to investigate alternative optimization techniques.

References

1. Selye, H.: The stress syndrome. *The American Journal of Nursing* 65, 97–99 (1965)
2. Hoffman-Goetz, L., Pedersen, B.K.: Exercise and the immune system: a model of the stress response. *Immunology Today* 15, 382–387 (1994)
3. The-American-Institute-of-Stress. America's No. 1 Health Problem - Why is there more stress today? (August 05, 2010), <http://www.stress.org/americas.htm>

4. Lifeline-Australia (Stress Costs Taxpayer \$300K Every Day (2009), <http://www.lifeline.org.au>)
5. Liao, W., et al.: A real-time human stress monitoring system using dynamic bayesian network. In: Computer Vision and Pattern Recognition - Workshops. CVPR Workshops (2005)
6. Zhai, J., Barreto, A.: Stress recognition using non-invasive technology. In: Proceedings of the 19th International Florida Artificial Intelligence Research Society Conference FLAIRS, pp. 395–400 (2006)
7. Dou, Q.: An SVM ranking approach to stress assignment. University of Alberta (2009)
8. Labbé, E., et al.: Coping with stress: the effectiveness of different types of music. *Applied Psychophysiology and Biofeedback* 32, 163–168 (2007)
9. Healey, J.A., Picard, R.W.: Detecting stress during real-world driving tasks using physiological sensors. *IEEE Transactions on Intelligent Transportation Systems* 6, 156–166 (2005)
10. Scherer, S., et al.: Emotion recognition from speech: Stress experiment. In: Proceedings of the 6th International Language Resources and Evaluation (LREC 2008), Marrakech, Morocco (2008)
11. Sharma, N., Gedeon, T.: Stress Classification for Gender Bias in Reading. In: Lu, B.-L., Zhang, L., Kwok, J. (eds.) *ICONIP 2011, Part III. LNCS*, vol. 7064, pp. 348–355. Springer, Heidelberg (2011)
12. Goldberg, D.E.: *Genetic algorithms in search, optimization, and machine learning*. Addison-Wesley (1989)
13. Park, B.J., et al.: Feature selection on multi-physiological signals for emotion recognition. In: 2011 International Conference on Engineering and Industries (ICEI), Korea, pp. 1–6 (2011)
14. Niu, X., et al.: Research on genetic algorithm based on emotion recognition using physiological signals. In: 2011 International Conference on Computational Problem-Solving (ICCP), pp. 614–618 (2011)
15. Yacci, P.: Feature selection of microarray data using genetic algorithms and artificial neural networks. Rochester Institute of Technology (2009)
16. Golub, T.R., et al.: Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 286, 531–537 (1999)
17. Ferreira, P., et al.: License to chill!: how to empower users to cope with stress. In: Proceedings of the 5th Nordic Conference on Human-Computer Interaction: Building Bridges, pp. 123–132 (2008)
18. Dishman, R.K., et al.: Heart rate variability, trait anxiety, and perceived stress among physically fit men and women. *International Journal of Psychophysiology* 37, 121–133 (2000)