

# Statistical Analysis on Human Factor Evaluation – Case Study using Artificial Neural Networks

<sup>1</sup>S.Subhashini, <sup>2</sup>Dr.R.Muruganandham, <sup>3</sup>Dr.K.Ravichandran, <sup>4</sup>S.Senthilvel, <sup>5</sup>E.Manokaran

<sup>1,5</sup>Dept. of Industrial Engg, TCE, Madurai, India

<sup>2</sup>Dept. of Mechanical Engineering, TCE, Madurai, India

<sup>3</sup>Dept. of Management Studies, Madurai Kamaraj University, Madurai, India

<sup>4</sup>Thiagarajar College of Engineering, Madurai, India

## Abstract

In current corporate scenario, the performance of human beings has become more challenging due to several aspects like work environment, peers competition, organizational conflicts, working stress, project deadlines, etc. . Among all factors stress plays a major role in motivating or de-motivating an individual. Our paper has tried to peruse the origin, nature and effects of stress for budding professionals like engineering students. The paper has categorized the factors responsible for causing the stress to the professional students of Information Technology. Optimization technique ANN (1) tool has been used to identify the stressed professionals. The results have been validated using Statistical Analysis and determine the most significant factor for causing stress.

## Keywords

Multi-criteria Decision Making (MCDM), Artificial Neural Network (ANN), Optimization Technique, Statistical Analysis, Feed Forward algorithm

## I. Introduction

Optimization is an active and fast growing research area and has a great impact on the real world. Despite of the enormous amount of work that has been conducted both theoretically and empirically and the huge success that has been achieved in different aspects, it is still an ongoing and long-term task to develop competent techniques, which could effectively solve large-scale optimization problems. Optimization technique is one of the popular methods to find a solution for complex decision-making problem and also finding ranking performance under the given constraints, by maximizing desired factors and minimizing undesired ones. In this thesis, ANN methodological tool have been used to validate the stress causing factors and to rank according to the scores and validate the results using statistical analysis to determine the most significant factor.

## II. Artificial Neural Network (ANN)

An Artificial Neural Network (ANN)[1] is an information processing paradigm that is inspired by the way biological nervous systems, such as the brain, process information. The key element of this paradigm is the novel structure of the information processing system. It is composed of a large number of highly interconnected processing elements (neurons) working in unison to solve specific problems. An ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process. Learning in biological systems involves adjustments to the synaptic connections that exist between the neurons, and for the validation process ANN is followed, the human brain provides proof of the existence of massive neural networks that can succeed at those cognitive, perceptual, and control tasks in which humans are successful.

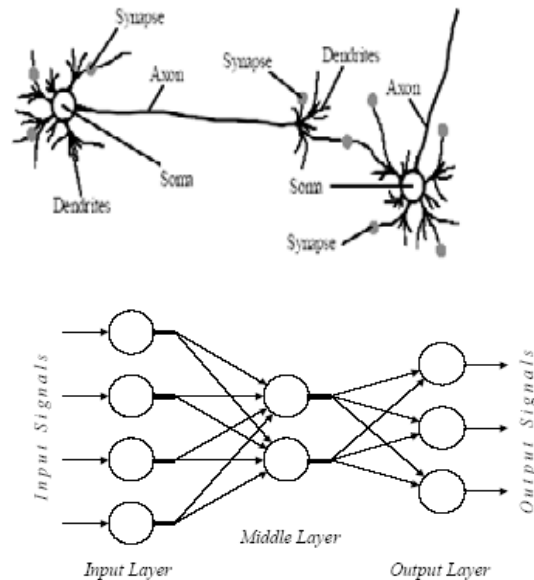


Fig. 1:

## III. DATA COLLECTION

For the analysis of students stress level[5] the data has been collected among various students by conducting pilot study, questionnaire technique. Based on the data's ranking of each student can be determined using Back propagation algorithm [5]

Table 1 :

Student No	Institutional Factor Score	Personal factor Score	Environmental Factor score
1	27	38	27
2	16	20	10
3	30	40	28
4	28	39	26
5	12	29	20
6	17	32	23
7	23	31	21
8	26	30	25
9	18	32	24
10	25	31	23
11	22	34	23
12	15	33	23
13	13	32	24
14	15	31	26
15	13	33	19
16	26	21	16
17	23	33	25

18	20	30	21
19	17	31	23
20	23	31	21
21	16	24	14
22	25	35	25
23	22	34	22
24	15	33	23
25	13	32	21
26	26	23	13
27	23	31	21
28	20	30	21
29	12	28	20
30	17	32	23

#### IV. Back propagation algorithm

Back propagation [5] is a systematic method of training multi-layer artificial neural network. A Back Propagation network consists of at least three layers of units: an input layer, at least one intermediate hidden layer, and an output layer. In BP units are connected in a feed-forward fashion with input units fully connected to units in the hidden layer and hidden units fully connected to units in the output layer.

- $Y = f(w.p + b)$

$P = (P_1 \dots P_r)$  is the input column vector.

$W = (w_1 \dots w_r)$  is the weight vector.

- Perceptron Learning rule:

If  $e = 1$ , then  $w_{new} = w_{old} + p$ ,

$B_{new} = b_{old} + 1$ .

Based on stress value we have to analysis what are all the factors for causing stress. Among various factors I consider only four major factors for IT department students. They are FAT, BMI; Time spends in system, Physical activity.

Input data for each factor: For this input data find weight age value for each factor using back propagation algorithm,

Sample Calculation for Weight age value:

Average value for each factor,

FAT = 699.5

BMI = 664

Time spends in system = 178

Physical activity = 370.

Based on this value weight age value for each factor is,

$W_1 = 1, w_2 = 0.949, w_3 = 0.254, w_4 = 0.528$

$P = ((35 * 1) + (23 * 0.949) + (8 * 0.254) + 0) + 1$

Where  $b = 1$

$P = 58.8 + 1 = 59.8 * 0.1 = 5.98$

$W_{new} = w_{old} + p$

$W_{1new} = 6.98, w_{2new} = 6.92, w_{3new} = 6.23, w_{4new} = 6.5.$

Hence this weight age values are consider as an input values for doing Statistical analysis.

Table 2 :

Rank	Student no	Stress value %	Fat%	BMI in kg/m <sup>2</sup>	System time spends in hours per day	Physical activity in min per day
1	42	83.56	35	23	8	0
2	3	81.19	32	22	7	0
3	4	77.1	28	21	8	10
4	1	76.2	27	23	8	0
5	22	70.46	27	20	7	0
6	38	70.45	26	22	8	0
7	32	69.58	25	23	4	0
8	17	67.15	23	21	8	15
9	8	67.04	22	22	8	10
10	35	66.43	21	21	3	0
11	11	65.56	22	22	4	0
12	10	65.45	26	27	4	0
13	31	64.74	24	26	3	0
14	23	64.74	23	21	8	20
15	34	63.82	22	24	8	15
16	7	62.2	20	23	7	0
17	20	62.2	21	21	8	15
18	27	62.2	23	23	3	10
19	33	61.15	20	20	8	10
20	9	61.44	22	22	7	0
21	39	60.73	23	23	8	0
22	6	59.81	20	21	3	0
23	30	59.81	20	20	7	0
24	14	59.79	21	23	3	0
25	12	59.65	21.5	23	3	0
26	24	59.04	22	21	8	20
27	36	59.04	20	21	8	15
28	19	58.92	24	22	3	20
29	18	58.92	20	23	3	180
30	28	58.18	19	20	3	30

#### V. Analysis Of Results

Statistical analysis [7] gives us a way to quantify the confidence we can have in our inferences. It helps us to analyze various factors in graphical manner. Using MS-Excel ANN results can be validating. In statistical analysis first correlation test has been taken between factors and students stress value to find out which factor has more relation.

#### A. Correlation

Correlation is a statistical measurement of the relationship between two variables. Possible correlations range from +1 to -1.

$$R(x,y) = \text{cov}(x,y) / \sigma_x * \sigma_y.$$

$$\text{Cov}(x,y) = 1/n \sum xy - \bar{x} \bar{y}.$$

$$\sigma_x^2 = 1/n \sum x^2 - \bar{x}^2$$

$$s_y^2 = 1/n \sum y^2 - \bar{y}^2$$

Using this formula the values are found,

FAT = 0.93

BMI = -0.04

Physical activity = -0.3

Time spends in system = 0.41.

From above values it seems the factor FAT has more relation with stress.

**1. ANOVA**

ANOVA stands for analysis-of-variance, a statistical model meant to analyze data. Generally the variables in an ANOVA analysis are categorical, not continuous.

Hence it confirmed the factor FAT has more effect on stress level. FAT has various ranges on human body. To determine which range has more adversely affected find ANOVA and compare the results with residuals.

NULL HYPOTHESIS: Factor FAT has no significant effect on stress level

ALTERNATE HYPOTHESIS: FAT has significant effect on stress level

In ANOVA formula for determining significant range is,

FAT	STRESS					TOTAL Y <sub>i</sub> .		AVG
15	62.2	61.1	59.8	59.0	58.9	58.1	359.1	59.8
20	64.71	63.8	62.2	61.4	60.7	59.04	371.8	61.9
25	69.5	67.1	65.4	64.7	62.2	58.9	387.8	64.6
30	77.1	76.2	73.4	70.4	75.4	73.2	445.7	74.2
35	83.5	81.1	79.8	81.25	82.1	80.2	487.9	81.31
Y <sub>j</sub> .	357.01	349.3	340.6	336.7	339.3	329.4	Y <sub>..</sub> = 2052.3	
AVG	71.40	69.8	68.1	67.3	67.8	65.8	Avg = 410.46	

1.  $SS_{treat} = 1/n \sum y_{i..}^2 - Y_{..}^2/N$

2.  $SS_{total} = \sum y_{ij}^2 - Y_{..}^2/N$

3.  $SS_{error} = SS_{total} - SS_{treat}$

Using above formula the values

Sum of squares	DOF	Mean squares	F <sub>o</sub>
SS treat	4	496.72	87.9
SS total	29		
SS error	25	5.65	

Compare with table value at  $\alpha = 0.05$ , F 4, 25 = 2.75

F<sub>o</sub> > F<sub>α</sub>

Therefore we reject the H<sub>0</sub>.

**2. Residuals**

In statistics and optimization, statistical errors and residuals are two closely related and easily confused measures of the deviation of a sample from its "theoretical value".

Residual =  $y - \hat{y}$

Find the range which has more effect for causing stress.

$E_{ij} = y_{ij} - \hat{y}_{ij}$

$\hat{Y}_{ij} = y_{i.} + y_{.j} - y_{..}$

Using this formula find the values and plot it.

Table 3 : Residuals

FAT	E11	E21	E31	E41	E51	E61

15	341.46	341.96	342.36	342.36	341.76	342.96
20	341.8	342.4	339.9	342.5	341.3	341.7
25	343.9	343.1	343.1	343.2	330.2	338.9
30	341.9	342.6	341.5	339.3	343.8	343.6
35	341.2	340.4	340.8	343.1	343.4	343.5

From the above table the residuals E51 have more variation.

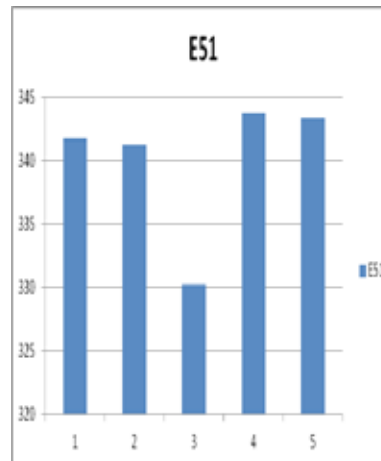


Fig. 5.1

Hence finally it conclude if the FAT range is more than 30% in human body then the stress level also get increased.

**3. Probability plotting**

Probability values are found using the formula,

$P = j - 0.5 / n$ , Using z- table values are found for each student.

Normal probability graph is drawn between residuals and probability value to find interaction between factor and stress

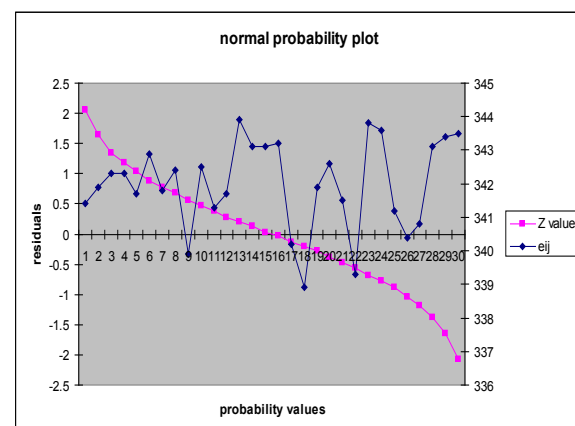


Fig. 5.2

Normal probability graph is drawn between residuals and probability value to find interaction between factor and stress

**VII. Conclusion**

It is quite clear that selection of students with human factor involves a large number of considerations. The use of

statistics and ANN method is observed to be quite capable and computationally easy to evaluate and select significant effect of stress from a given data. Thus stress plays a major role in human life. Hence causes for stress in students life can be predict and analyze is a complex one for that here I use ANN and statistical tool for analyzing. Ranking of students stress level is done using neural network. Analysis of factor effect for a particular department is made. From the analysis it shows the factor FAT has more effect on stress in IT department students

### References

- [1] Ajith Abraham Oklahoma the Book, "Artificial Neural Networks", State University, Stillwater, OK, USA.
- [2] Bersam Bolat, Gül Tekin Temur, Umut Gündüz, Metin Yılmaz, "Evaluation of Managers' Perceptions on Information System Project Success using ANN compared to MDA: A study of Retailing Industry", European and Mediterranean Conference on Information Systems 2009.
- [3] Coryn A.L. Bailer-Jones Max-Planck-Institute, "An introduction to artificial neural networks", für Astronomie, K"onigstuhl Vol. 17, Germany.
- [4] Laura P. Womble, "Impact of Stress factors on college students Academic Performance", University of North Carolina at Charlotte, 2003.
- [5] M.Lilly Florence, P. Balasubramanie, "Multilayer Feed Backward neural Network Model for Medical Decision support: Implementation of Back Propagation Algorithm in HIV/AIDS Regimens", International Journal of Reviews in Computing 2009 IJRIC.
- [6] Sunita Yadav, "Neural Network based Approach for Predicting User Satisfaction with Search Engine", International Journal of Computer Applications Volume 18- No.5, 2011.
- [7] Warren S. Sarle, "Neural Networks and Statistical Models" Proceedings of the Nineteenth Annual SAS Users Group International Conference, April, 1994.
- [8] Adele Alijanpour shalmani, Mahmoud Shabanpour Shahrestani, Hossein Asadi and Farid Bagheri, "Comparison of regression pedotransfer functions and artificial neural networks for soil aggregate stability simulation", World Congress of Soil Science, Soil Solutions for a Changing World 1 - 6 August 2010, Brisbane, Australia.



S. Subhashini received her B.E. degree in Electronics And Instrumentation Engineering from Kamaraj College of engineering and Technology, Virudhunagar, Madurai, India in 2010. The M.E. degree in Industrial Engineering from Thiagarajar College of Engineering, Madurai, India, in 2010-2012. Her research interests include Optimization

techniques.



Dr. R. Muruganandham currently working as a Assistant professor in Thiagarajar college of Engineering, Madurai, India.



Dr. K. Ravichandran is currently working as a Professor in Management Studies, Madurai Kamaraj University, Madurai.



S. Senthilvel currently doing his B.E. degree in Electronics And Communication Engineering from Thiagarajar college of Engineering, Madurai, India.



E. Manokaran received his B.E. degree in Electronics and Communication Engineering from Adhiaman College of engineering, Hosur, India in 2010. Doing M.E. degree in Industrial Engineering from Thiagarajar College of Engineering, Madurai, India, in 2010-2012. His research interests include Multi criteria decision making.