IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-6, Issue-6, June 2018 E-ISSN: 2347-2693

An Investigation of Occupational stress Classification by using Machine Learning Techniques

S.K. Yadav[1], Arshad Hashmi[2*]

[1] Dept. of Computer Science, Shri Jagdishprasad Jhabarmal Tibrewala University Jhunjhunu, Churu Road, Vidyanagari, India
[2] Dept. of Computer Science, Shri Jagdishprasad Jhabarmal Tibrewala University Jhunjhunu, Churu Road, Vidyanagari, India

*Corresponding Author: hashmi.arshad80@gmail.com

Available online at: www.ijcseonline.org

Accepted: 19/Jun/2018, Published: 30/Jun/2018

Abstract. Occupational stress can impact our lives in several aspects. This affects employee's health, causes absenteeism and overall performance of an organization affected. World Health Organization (WHO) identifies it as epidemics for the modern life. The insurance sector employees have direct customer interaction. The policies and the services introduced to the new customers, convincing the ideas and satisfying the divergent customer needs causes more pressure on the employees which leads to higher level of stress. Occupational stress data mining is an emerging stream which helps in mining stressed data for solving various types of problem. One of the problems is to know the impact of role overload and role ambiguity on occupational stress. In this paper, we have tried to implement a model using machine learning classification techniques for the prediction of Occupational stress related to insurance sector personnel. In this paper, we have applied support vector machine (SVM), Neural network (NN), decision tree (DT) and random forest (RF). The training and testing are done through a stratified tenfold cross-validation. The proposed model obtained an accuracy of 60%, a sensitivity of 80%, and specificity 60%. The evaluation of occupational stress is critically connected to job performance in the organization. So it is essential to identify the causes of occupational stress and can be reduced to the possible extent with the help of proper management techniques.

Keywords. Occupational Stress, Distress, Predictive model, Classification techniques, SVM, NN, DT, RF

I. INTRODUCTION

The number of peoples suffering from a stress is increasing day by day in modern society. Stress and overall health condition are interlinked. In terms of Hans Selye stress is "a non-specific response of an organism in exposure to a demand or a change in the physical situation [1]". It appears either in the form of eustress or distress. Eustress has the positive impact while distress has the negative impact, an issue of concern for the employee. A pattern developed by means of cognitive, emotional, behavioral and physiological reactions to adverse and noxious aspects of work content, work organization and work environment [2] comes under occupational stress. It is the result of poor matches between people and their related work and conflict between their parts at work and the work environment.

"Occupational stress" is a comparatively modern subject, which got focus in recent days as a major health-related concern [3, 4] based on the National Institute for Occupational Safety and Health (NIOSH). "Occupational stress" causes serious concern for operating any organization effectively.

In the Insurance sector, employees have to introduce as well convince the investors related to insurance products and policy, which causing stress-inducing situations. During the discussion of the insurance policy details with the future investors, it is in the best interest to manage the stress levels in order to obtain a favorable outcome. The main reason behind the heavy stress of the insurance sector employees includes role conflict, role overload, role ambiguity, organizational culture, long working hours, lack of support from management to the employees. The insurance sector employees notice a good number of symptoms so they are under a high level of stress because of their work. To identify the potential reasons for negative stress to undermine one or more number of workers there is an availability of occupational stress index and also various other tools. This will help employees as well as the employer to know the correct vision for the workplace with objectively. Self-reported measurements are the most common way to asses stress from a psychological standpoint. It has been found from the literature review that data mining tools have not been yet utilized commonly to analyze occupational stress issues [5, 6, 7]. In order to analyze datasets for obtaining useful knowledge data mining techniques have been utilized [8, 9, 10].

1.1 Research Problem

After reviewing the literature it is found very less work has been done so far for evaluating occupational stress of employee specifically to the domain of insurance sector personal by using data mining techniques. This fact gives an idea to do this study for evaluating the pervasiveness and linked parameters of occupational stress among the employee of the insurance sector. In most of the studies to discuss job stress evaluation issues, classification of job stress factors has been performed by focusing on psychosocial and other environmental factors without taking care of physical factors. For example [11] underlined psychosocial factors in past job contents questionnaire (JCQ). To this end, we propose a survey method called the Occupational Stress Index (OSI) developed by Srivastava and Singh. This survey incorporates all three factors psychosocial, environmental and physical factors. This OSI measures the extent of job-related stress related to public and private life insurance sector employees.

1.2 Objective of the study

The objectives of the present study are

1. To investigate the impact of role overload and role ambiguity on occupational stress using machine learning classification techniques with special reference to the employee of Insurance sector
2. To suggest a best occupational stress prediction model.
3. To analyze the various components of stress and identify the stress management for public and private sector of insurance industry.
4. To give suggestion to overcome stress in insurance sector.

The process of detecting stress using physiological self-assessment in the form of questionnaires consists of following phases. See Figure 1.

First, data are collected from questionnaires. Secondly, cleaning the data and statistical analysis performed to check the validity of data. Finally on the same dataset several machine learning techniques applied by means of R coding. The results are compared to obtain the best performing model.

The remaining section of this paper is organized in the following way; Section 2 has brief literature review. The data mining techniques used in the experiments and the information about the dataset are presented in Section 3. Section 4 has experiments and discussion. The paper ends with a conclusion.

II. RELATED WORK

Various studies are carried out to study the occupational stress and its related factors. Krishna et al. [12] have applied Analysis of variance including classification and regression tree (CART) for measurement and modeling of Job Stress. To explore the Effective Factors on Job Stress Arezou Khaleghi et al. [13] have applied correlation coefficient test and progressive multivariate regression. To evaluate stress arising from psychosocial risk factors multiple regression analysis is used by Serpil Aytac [14]. Mohd Zuri Ghani et al. [15] discuss the relationship between the job stressor and health by using demographic variables gender, age, marital status, highest academic qualification, job tenure, job title, and hierarchical level. Che Noriah Othman et al. [16] uses Occupational Stress Index of Malaysian University workplace to evaluate the level of stress experienced by Government University staffs in Malaysia. Chi-Square Automatic Interaction Detection (CHAID) decision tree algorithms have been used by Shaghayegh Parhizi et al. [17] to explain the relationship between fatigue dimensions and psychosocial factors among registered nurses. Maryam Khodabakhshi [18] found in her study that women bank employees with an assessment of their organizational commitment and personality type experienced significantly higher levels of occupational stress. Susana García-Herrero et al. [19] propose a probabilistic model using Bayesian networks to explain the relationships between work demands and occupational stress. For identification of stress, Yong Deng et al. [20] explains how feature selection method is used in combination with Principal Component Analysis (PCA). For individual stress diagnosis, Shahina Begum et al. [21] have proposed Case-based decision support system based on the use of Fuzzy logic in combination with case-based reasoning (CBR). Further Arie Shiro et al. [22] have done an analysis of how three socio-demographic variables, employee gender, age, and tenure are interrelated with job performance by using Weighted Least Squares (WLS) regression. Eric G. Lambert et al. [23] discuss job stress by proposing Multivariate models and they employed five separate multivariates. Jing Zhai, et al. [24] elaborate Realization of Stress Detection using Psychophysiological Signals for Improvement of Human-Computer Interactions STROOP TEST, Support Vector Machines (SVMs), Linear, RBF and Sigmoid Kernel. Vedat Isikhan et al. [25] in his study explain the Job stress and coping strategies for healthcare professionals working with cancer patients by means of T-test and SPSS. Noora Nenonen [26] perform an Occupational accident Data mining using decision tree and association rules. Jorn Bakker et al. [27] discuss stress at work by means of associative classification in data mining. Ming Jiang et al. [28] have discussed a method for Stress Detection Based on fuzzy c-means (FCM) Clustering Algorithm. Paul Bowen et al. [29] discuss in his study Occupational stress and job demand, control and support factors using Hierarchical regression, factor analysis and structural equation modeling (SEM). It is found from the review that in most of the cases traditional statistical analysis, such as regression models, analysis of variance

(ANOVA), and hierarchical linear models have been extensively used. In some cases, data mining techniques have also been used to study occupational stress. In the next section, we are going to discuss the methods which we have to follow during the research work.

III. METHODOLOGY

Proposed integrated procedure for evaluating job stress

During this study, we will discuss the matter in the following way. The first step is to collect the data by means of a survey using OSI Questionnaire by incorporating potential occupational stress reasons. Then we compile and analyze the data after performing preprocessing task. Then we perform statistical analysis to decide classification task so that data can be transformed for application of machine learning classification techniques. We will apply filtering and feature selection before training the data by applying several machine learning classification techniques. By analyzing the comparison of different classifier best prediction model will be chosen. An overview of the proposed method is illustrated in Fig. 1

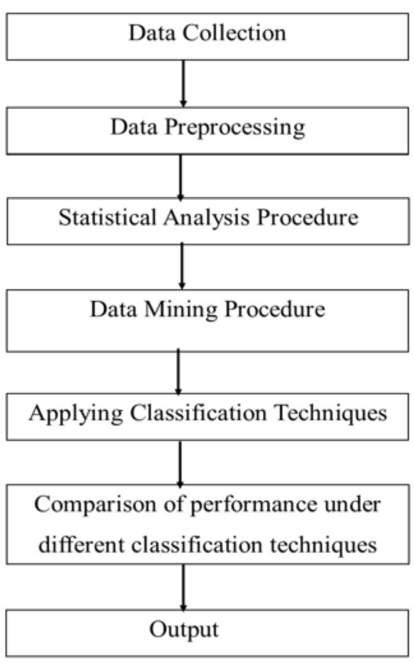

Figure.1: Overview of the Procedure for Evaluating Job Stress

3.1 Occupational Stress Questionnaire

The modified Occupational Stress Index is used for data collection. Demographic variables are used to measure the personal information of the respondents. This OSI examines occupational stress in twelve particular dimensions. These dimensions are role overload, role ambiguity, role conflict, unreasonable group and political pressure, responsibility for persons, under participation, powerlessness, poor peer relations, intrinsic impoverishment, low status, strenuous working conditions and unprofitability. The questionnaire contains 46 statements which measure twelve types of variables for the study. These 46 items are related with twelve particular dimensions. The purpose of these twelve variables and 46 statements is to find out the actual stress level of an employee. The questionnaire is based on Likert's five-point scale [34] and each of 46 statements has five options of answers. The scale options are given: viz., 'Strongly Agree'; 'Agree'; 'Neutral'; 'Disagree' and 'Strongly Disagree'.

This OSI measures the extent of job-related stress that employees of both public and private life insurance sector perceived as arising from various constituents and conditions of their jobs. This scale has been found to have high reliability and has proved its validity through experiments; therefore, it has been used for this study. In this paper, we are trying to investigate the impact of Role overload and Role ambiguity for job stress.

The Role Overload is related with work-load, lack of staff, job dissatisfaction etc. while Role Ambiguity is related with insufficient information about the assignments and planning of job, ambiguous expectations by colleagues and supervisors, etc.

Work Stress questionnaire consistency measured by Cronbach's [35].The obtained value is found 0.751 which is above than the 0.7, the accepted threshold value. In the next section, proposed data mining techniques that are used in the experiments are discussed

3.2 Data mining (DM) techniques for the analysis

In this paper, we are going to investigate the impact of role overload and role ambiguity on the occupational stress using classification techniques [30] and comparing the performance of different techniques.

3.2.1 Decision Trees- Decision trees are very popular classification techniques because of their easy interpretability [31]. These trees are built by using divide and conquer method. The building process starts with a root node, the data points at each label is divided on an attribute selected by the chosen split criterion. The tree building process stops by the selected stopping criterion. The path between the root node to a leave acts as a rule to predict the decision. The information in leaves is the predicted values. In this paper Rep Trees and J48 [32] are used for experiments.

3.2.2 Support Vector Machines (SVM) -. It divides a line to give the best separation of the data into the two groups by means of an optimization process. The data instances that are closest to the line are considered during this process. These instances are called support vectors. Kernels play an important role to control the projection in order to make the separation of class flexible. Hence selecting a proper kernel is an important step in SVM.

Prediction will be done with the help of following equation

$$\mathbf{f(x) = B0 +} \sum_{i=1}^{n}\left((a_i) \times (x \times x_i)\right)$$

x_i – represents support vector

x- represents new input vector

B0 and a_i - co-efficient for each input that are calculated from the training data by means of learning algorithm.

3.2.3 Artificial Neural Networks (ANN)

This type of network used a biological neural network that has the characteristics to learn and react. It became a popular technique and can be used in classification problems related to health. It is a new approach to stress research. Stress models using this type of network are at early stage of research and developed promising results. It is found that this network gives better result to recognize stress as compared to humans beings for given voice recordings [Scherer, 2008]. These results can contribute more motivation in the area of stress research. It learns patterns to recognize characteristics from input tuples. It consists of processing elements which are interconnected together known as artificial neurons. These processing elements are connected by weighted links which can pass the signal between neurons. It is inspired by the biological neural network and it learns from its inputs. It can approximate the functions and classify the patterns. This feed-forward neural network is known by its simplicity shown in Figure 2.4. This network contains three layers i.e Input layer, Hidden layer, and the output layer. The input layer is used to present patterns and at hidden layers weights of neurons are adjusted which can communicate further to the output layer. At output layer evaluation of all the outputs is done. In ANN, a signal is propagated layer after layer. This topology is designed in such a way that includes a time-delay to overcome the time-varying nature of the stress. To determine its topology, it is important to find the number of points in stress classification. This topology has an effect on the performance for classification.

IV. RESULTS AND DISCUSSION

The experiments are performed using R software [45]. R is open source software which has many classification methods implemented. The default values of different parameters are used for experiments 10 fold cross-validation strategy is used in the experiments.

4.1 **Classification**

The classifier algorithm established the relationship between a set of correctly labeled objects and unknown objects in order to classify the unknown objects [46] correctly. Using comparison of metrics, obtained with a set of classifier algorithms, best performance model [47] is selected, with the help of Support Vector Machine (SVM) [49] with the polynomic function [50]; J48 as well as Random Forest and neural network. A decision tree model is a good preliminary exploratory technique.

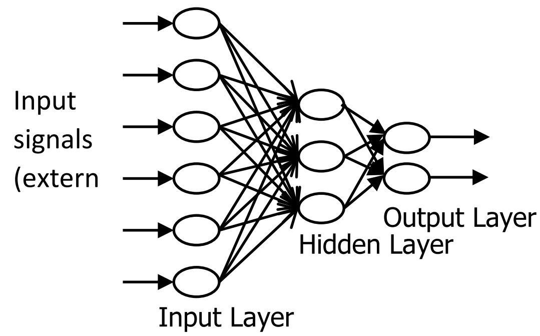

Figure 2 Multilayer feed forward network

The standard evaluation technique "a stratified 10-fold cross-validation" is used for the available limited data[52]. Performance evaluated by means of accuracy which is related to the percentage correctly classified among all employees. Sensitivity (sens) which is related with the percentage correctly classified as stressed among all classified as stressed. While the specificity (spec) is related with percentage correctly classified as stressed among all really stressed. The ROC sure represents the combined measure between true positive(TP) and False Positive(FP). The performance results are shown in corresponding graphs

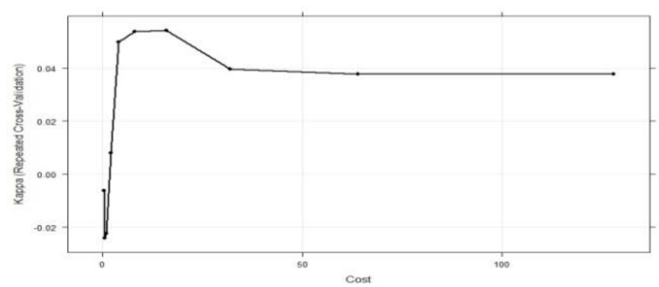

Figure 4.1 a

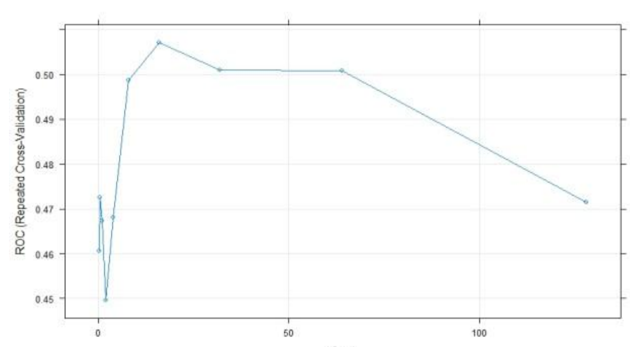

Figure 4.1 b

Figure 4.1 ROC and Accuracy on different value of parameters (decision tree)

Observation1: In above figure 4.1 show decision tree performance on Accuracy and ROC (receiver operating curve) is the ratio between sensitivity and specificity. This experiment performs on R language with 55 features and two classes high stress and low stress(normal).In the experiment using 500 instances and ten cross-validation and iterative run on 0 to 1 different kappa parameter. Kappa parameter shows the complexity of the model in both graph(a) and (b) y-axis show the kappa parameter cost and x-axis accuracy and ROC value. Maximum accuracy goes with 55 features is 52% in case of a decision tree and ROC goes to 64. this means sensitivity high than specificity.

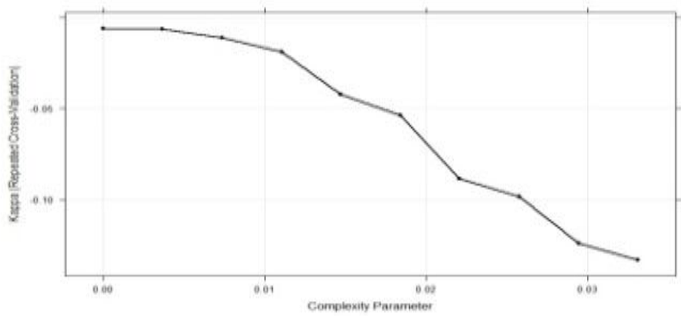

Figure 4.2 a

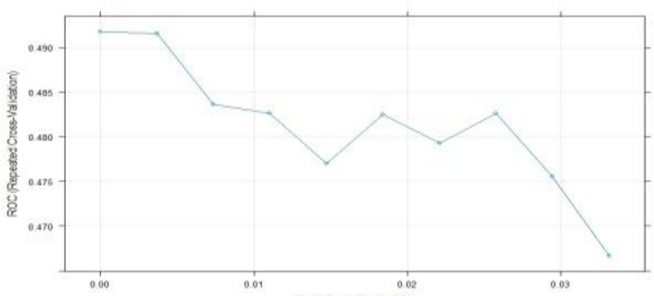

Figure 4.2 b

Figure 4.2 Show the ROC and Accuracy of SVM on different Value of Parameters

Observation2: In above figure 4.2 show decision tree performance on Accuracy and ROC (receiver operating curve) is the ratio between sensitivity and specificity. This experiment performs on R language with 55 features and two classes high stress and low stress(normal).In the experiment using 500 instances and ten cross-validation and iterative run on 0 to 1 different complexity parameter. Kappa parameter shows the complexity of the model in both graph(a) and (b) y-axis show the complexity parameters(C) and x-axis accuracy and ROC value. Maximum accuracy goes with 55 features is 53.4% in case of a decision tree and ROC goes to 66. this means sensitivity high than specificity.

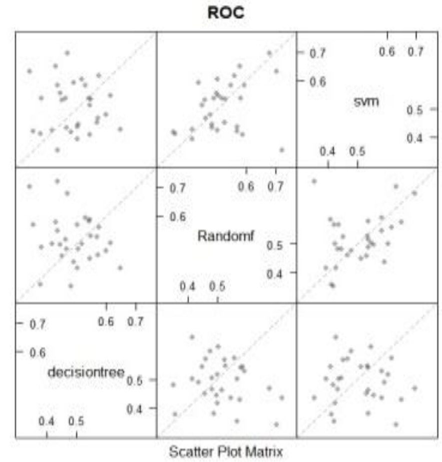

Figure 4.3 a

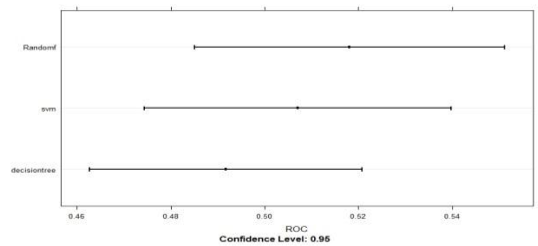

Figure 4.3 b

Figure 4.3 Comparative analysis of RF, SVM and decision tree of ROC and Accuracy

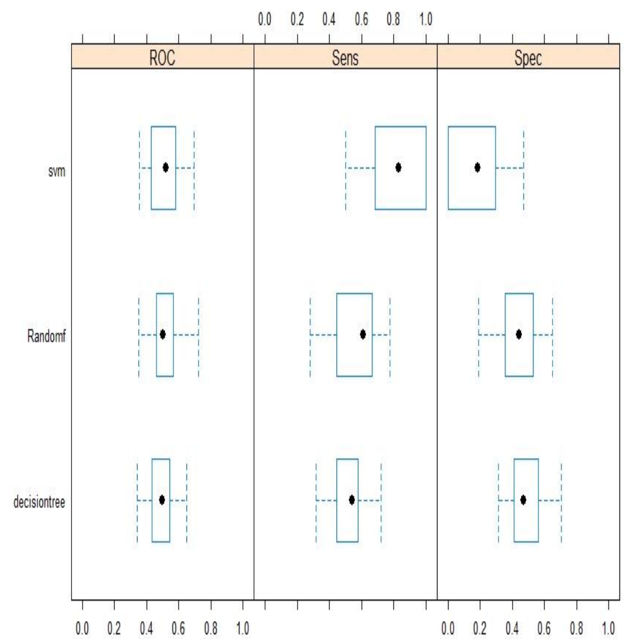

Figure 4.4 Comparative analysis of RF, SVM and decision tree of ROC, sensitivity and specificity

Observation3: In above figure 4.3 show decision tree performance on Accuracy and ROC (receiver operating curve) is the ratio between sensitivity and specificity. This experiment performs on R language with 55 features and two classes high stress and low stress(normal). In the experiment using 500 instances and ten cross-validations. In this experiment comparison of ROC, sensitivity, specificity, and accuracy. First, start from SVM which show high sensitivity approx. 80% but less specificity 45%. It means to show the effective result because in this problem sensitivity should be because if any high stress classified as normal then it makes a problem. If analysis the accuracy from fig4.3(b) SVM accuracy middle of a decision tree and random forest. Next analysis of random forest it shows the same performance in case of sensitivity and specificity but high accuracy compare to SVM and decision tree so we can say random forest perform well than other classifiers .

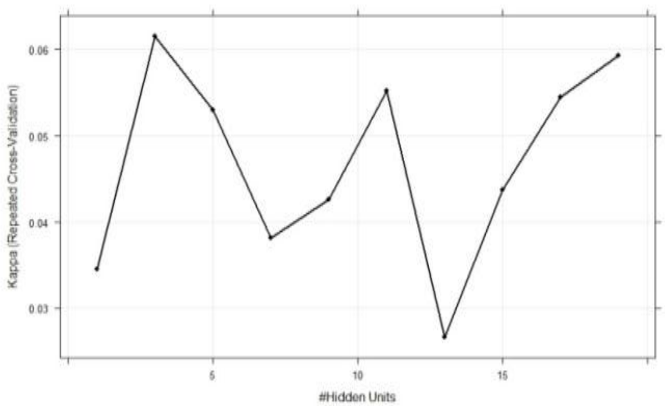

Figure 4.5 Neural network accuracy performance with increasing hidden units

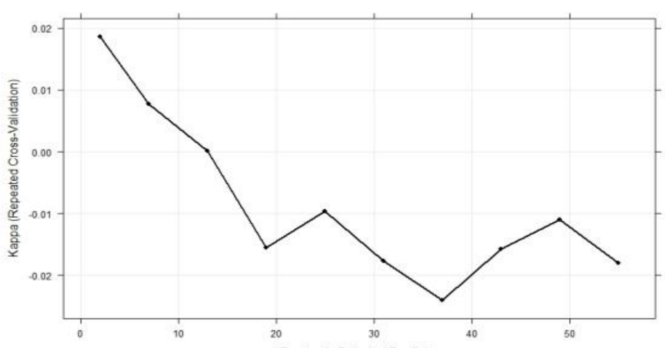

Figure 4.6 Random Forest accuracy performance with increasing Predictor units

Observation4: In above figure 4.5 and 4.6 increase the analysis on the resources of the classifier by hidden layers and number of predictors in neural network and random forest respectively. In both classifier random forest show shocking when increasing the predictor then accuracy reduces but neural network improves the accuracy performance 60% which encourage us to improve accuracy by a deep neural network. It is enhancing in future

V. CONCLUSION

The findings of a literature review study conducted by Dickinson and Wrigitt reported role overload as the major stressor [53]. This finding is consistent with that of a study conducted in Gorgan.[55].Occupational Stress is very common and it has a high cost in terms of employees' health, absenteeism, and lower performance. Occupational stress could be recognized through a continuous monitoring of employees during office work. In our experiment, we have evaluated the performance of different classifier and compared the result with the three different sets for investigating the effect of role overload on occupational stress. The same exercise has been done with three different data set for investigating the effect of role ambiguity on occupational stress. The best performing algorithm for role overload and role ambiguity is found to be Random Forest. In this paper analysis by metric (SVM, neural network) and non-metric (decision tree and Random forest).In which Metric approaches perform well in case of accuracy, sensitivity, specificity and ROC

VI. REFERENCES

[1] R. PJ. "The health effects of job stress". *Bus Health. PubMed, Vol.1Issue.6*, 198, pp.5–8,1984

[2] European Commission, "Guidance on Work-related Stress", Health and safety at work, Luxembourg, Directorate-General for
Employment and Social Affairs , Unit D.6, 1999

[3] WHO, "Raising awareness of stress at work in developing countries, a modern hazard in a traditional working environment,
Protecting workers",H*ealth series; no. 6*. 2007

[4] Allison G., "*Perceived and Reported Occupational stressors and Coping Strategies of Selected Community College Business* Faculty Members in Texas", University of Texas, United States; 2004 PhD dissertation.

[5] Bevilacqua, M., Ciarapica, F.E., Giacchetta, G, "*Industrial and Occupational Ergonomics in The Petrochemical Process
Industry: A Regression Trees Approach*", *Accid. Anal. Prev. Vol.40*, pp.1468-1479, 2008.

[6] Parhizi, S., Shahrabi, J., Pariazar, M., "*A New Accident Investigation Approach Based On Data Mining Techniques*", *J. Appl. Science, Vol.9*, pp. 731-737, 2009,

[7] Persona, A., Battini, D., Faccio, M., Bevilacqua, M., Ciarapica, F.E., "*Classification Of Occupational Injury Cases Using The Regression Tree Approach*". *Int. J. Reliab. Qual. Saf. Eng.Vol.13*, pp.171-191, 2006.

[8] *Giudici, P., "Applied Data Mining": Statistical Methods for Business and Industry. Wiley*, New York, 2003

[9] *Han, J., Kamber, M. Data Mining: "Concepts and Techniques"*, Morgan Kaufmann, San Francisco. 2001

[10] *Hand, D., Mannila, H., Smyth, P., "Principles of Data Mining"*. MIT Press, Cambridge (Mass.). 2001

[11] Karasek, R., Gordon, G., Pietrokovsky, C., Rrese, M., Pieper, C., Schwartz, J., Fry, L., Schirer, D., "*Job Content Questionnaire: Questionnaire and User's Guide", University of Massachusetts, Lowell*, 1985

[12] Krishna OB, Jhareswar Maiti, Pradip K. Ray, Biswajit Samanta, Saptarshi Mandal, Sobhan Sarkar, "Measurement and Modeling of Job Stress of Electric Overhead Traveling Crane Operators", *Science Direct*, http://dx.doi.org/10.1016/j.shaw.2015.06.005, 2015

[13] Arezou Khaleghi, Maryam Omidi Najafabadi, Frahad Lashgarara, "*Effective Factors on Job Stress from Experts' Perception; a Case Study in Iranian Agriculture Engineering Organization*", International Journal of Review in Life Sciences, Vol.5,Issue1, pp.94-99, 2015.

[14] Aytac Serpil, *The Sources of Stress, The Symptoms of Stress and Anger Styles as a Psychosocial Risk at Occupational Health*

[15] *and Safety: A Case Study on Turkish Police Officers*. 6th International Conference on Applied Human Factors and Ergonomics (AHFE 2015) and the Affiliated Conferences. Procedia Manufacturing Journal, 3, 6421 – 6428, 2015

[16] Mohd Zuri Ghani, Aznan Che Ahmad, Suzana Ibrahim ,Stress among Special Education Teachers in Malaysia, *Science Direct, Procedia - Social and Behavioral Sciences 114*, 2014 , 4 – 13

[17] Che Noriah Othman, Roz Azinur Che Lamin, Nursyuhadah Othman, Occupational Stress Index of Malaysian University Workplace, Procedia - *Social and Behavioral Sciences 153* ,2014 , 700 – 710

[18] Shaghayegh Parhizi, Linsey M. Steege, Kalyan S. Pasupathy, Mining the relationships between psychosocial factors and fatigue dimensions among registered nurses, *International Journal of Industrial Ergonomics 43*, 2013, 82-90

[19] Maryam Khodabakhshi, Predicting occupational stress for women working in the bank with assessment of their organizational commitment and personality type, Procedia -Social and Behavioral Sciences,84,2013,1859–1863

[20] Susana Garcia-Herrero, M.A.Mariscal, J.M.Gutiérrez, DaleO.Ritzel, Using Bayesian networks to analyze occupational stress caused by work demands-Preventing stress through social support, *Science Direct, Accident Analysis and Prevention 57*, 2013, 114– 123

[21] Yong Deng, D. frank hsu, Zhonghai wu, Chao-hsien chu , Combining multiple sensor features for stress detection using

[22] combinatorial fusion, *journal of interconnection networks, 1250008*, 2012,17 pages, doi: 10.1142/s0219265912500089

[23] Shahina Begum, Mobyen Uddin Ahmed, Peter Funk, Ning Xiong, Bo von Schéele, A Case-Based Decision Support System for Individual Stress Diagnosis Using Fuzzy Similarity Matching, *Computational* Intelligence, *vol 25,nr 3*, 2009,p180-195

[24] Arie Shirom, Simona Shechter Gilboa, Yitzhak Fried and Cary L. Cooper, Gender, age and tenure as moderators of work-related stressors' relationships with job performance: A meta-analysis, *SAGE Publications, Human Relations, 61(10)*, 2008, 1371–1398, DOI: 10.1177/0018726708095708

[25] Eric G. Lambert, Nancy L. Hogan, Marie L. Griffin, The impact of distributive and procedural justice on correctional staff job stress, job satisfaction, and organizational commitment, *Science Direct, Journal of Criminal Justice,35*, 2007,,644–656

[26] Jing Zhai, C. Chin, Chao Li, Realization of stress detection using psychophysiological signals for improvement of human-computer interactions, *IEEE Xplore Conference,* 2005, DOI:10.1109/SECON.2005.1423280

[27] Vedat Isikhan, Turhan Comez, M. Zafer Danis, Job stress and coping strategies in health care professionals working with cancer patients European *Journal of Oncology Nursing Society,8*, 2004, 234–244

[28] Noora Nenonen, Analysing factors related to slipping, stumbling, and falling accidents at work: Application of data mining methods to Finnish occupational accidents and diseases statistics database, *Applied Ergonomics, 44*, 2013, 215-224

[29]] J. Bakker, L. Holenderski, R. Kocielnik, M. Pechenizkiy, N. Sidorova ,TU Eindhoven, Stess@work:

From Measuring Stress to its Understanding, Prediction and Handling with Personalized Coaching, *ACM, IHI12*, 2012, 28–30

[30] Ming Jiang, Zhelong Wang, A Method for Stress Detection Based on FCM Algorithm, *IEEE*,DOI: 10.1109/CISP.2009.5304150,2009

[31] Paul Bowen, Peter Edwards, Helen Lingard, Keith Cattell, Occupational stress and job demand, control and support factors among construction project consultants, *International Journal of Project Management*, 32, 2014, 1273–1284

[32] Frawley, W., Piatetsky, G.S., Matheus, C., Knowledge discovery in databases: an overview. AI Magazine Fall, 213-228, 1992]

[33] Quinlan J.R., Induction of decision trees. *Machine learning, Vol.1*, 81-106, 1986

[34] Quinlan, 1992 Quinlan J.R, Learning with continuous classes, *Proc. of Australian Joint Conf. on AI*, 343-348, World Scientific, Singapore. . 1992

[35] Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., Crisp-DM 1.0: Step-by-step data mining guide. http://www.crisp-dm.org/ Retrieved November, 2009.

[36] Ammo,T. & Friedman, H. H, Overall Evaluation Rating Scale: An Assessment, *International Journal of Marketing Research,43(3)*, 301-310, 2002

[37] Cronbach, L. J, Coefficient alpha and Internal Structure of Tests, *Psychometrika, 16 (3)*, 297-333, 1951

[38] Travis C.Tubrea, Judith M.Collins, Jackson and schuler(1985), Revisited-a meta-analysis of the relationships between role ambiguity, role conflict, and job performance, *Journal of Management,26(1)*,155-169,2000, https://doi.org/10.1016/S0149-2063(99)00035

[39] Gray Balau, Testing the Longitudinal Impact of Work variables and Performance Appraisal Satisfaction on Subsequent Overall job Satisfaction, *Human relation, 52(8)*, 2009

[40] Carly S.Bruck,Tammy D.Allen,and Paul E.Spector,The Relation between Work–Family Conflict and Job Satisfaction-A Finer-Grained Analysis,*Journal of vocational Behavior 60*, 336–353, 2002,doi:10.1006/jvbe.2001.1836,available online at http://www.idealibrary.com

[41] J. Han, M. Kamber, and J. Pei, *Data Mining. Concepts and techniques, Third. Morgan Kaufmann*, 2012

[42] M. A. Hall, Correlation-based Feature Selection for Machine Learning, *The University of Waikato*, 1999

[43]] R. Kohavi and G. H. John, Wrappers for feature subset selection, *Artif. Intell. 97(1, 2)*, 273–324, 1997

[44] I. H. Witten, E. Frank, and M. Hall, Data Mining: *Practical Machine Learning Tools and Techniques, 40(6)*, Third Edit., 2001

[45] N. V Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, SMOTE: Synthetic Minority

[46] Over-sampling Technique, *J. Artif. Intell. Res., 16*, 321–357, 2002

[47] A. More, Survey of resampling techniques for improving classification performance in unbalanced datasets, *10000, 1–7*, 2016, https://arxiv.org/pdf/1608.06048.pdf

[48] Hall M.A., Frank E., Holmes G., Pfahringer B., Reutemann, P. and Witten, I.H, The WEKA Data Mining Software: An Update, *SIGKDD Explorations. 11(1)*, 10-18.2009

[49] O. Loyola, M. A. Medina, and M. García, Inducing Decision Trees based on a Cluster Quality Index, *IEEE Lat. Am. Trans.,13(4)*, 1141–1147, 2015

[50] P. Univaso, J. M. Ale, and J. A. Gurlekian, Data mining applied to forensic speaker identification, *IEEE Lat. Am. Trans., 13(4)*, 1098–1111, 2015.

[51] G. H. John and P. Langley, Estimating continuous distributions in Bayesian classifiers. *UAI'95 Proc. Elev. Conf. Uncertain. Artif. Intell.* , 338–345, 1995

[52] J. C. Platt, Fast training of support vector machines using sequential minimal optimization, in Advances in Kernel Methods: Support Vector Learning, *C. J. C. Burges, and A. J.Smola, Eds. MIT Press*, 185–208, 1999.

[53] B. Schölkopf and A. J. Smola, *Learning with kernels: support vector machines, regularization, optimization*, MIT Press, 2002.

[54] D. W. Aha, D. Kibler, and M. K. Albert, Instance-based learning algorithms, *Mach. Learn., 6(1)*, 37–66,1991.

[55] I. H. Witten, E. Frank, and M. Hall, Data Mining: *Practical Machine Learning Tools and Techniques, Third Edit., 40(6)*, 2001.

[56] Dickinson T, Wrigiit M. Stress and burnout in forensic mental nursing: A literature review. Br J Nurs. PubMed 17, 82–7., 2008

[57] Habrani P, Behdad F, Mobtaker M. Investigation of stressful factors in nurses., J Psychiatric Health Principles.3:231–7. 2008

[58] Kolagari SH, Sanagou A, Mirkarimi F, Behnampour N. The level of stress among gorgan university of medical sciences hospital operating rooms personals and its

relation to some related factors. J Gorgan Univ Med Sci.10:54–9. 2002

Authors Profile

Mr. Arshad Hashmi received Post Graduate Diploma in Computer Science and Application from University of Patna, India in 2002 and Master of Information Technology in 2010 from Punjab Technical University,Chandigarh,India 2010. He also received Master of Physics with Special as Electronics from University of Patna, India in 2007.He is Cisco Certified Network Professional.Cuurently, he works as a Lecturer in Faculty of Computing and Information Technology in Rabigh (FCITR), King Abdulaziz University, Kingdom of Saudi Arabia, since 2011.He is a member of IEEE computer society since 2016. He has published several research papers in reputed international journals.. He has more than 7 years of teaching experience and 3 years of Research Experience.

Dr.S.K.Yadav received Ph.D degree in Mathematics in 2011 from the B.R. Ambedkar University Lucknow as well as M.Phil. in Computer Science in 2007 from M.K.U.Madurai, India. He is associated as an Academics and Research Director in Shri JJT University, Churu, Rajasthan.In addition to this he is also assoicated with several other Academics and Research institutions.

His main research work focuses on Mathematical Modeling and Graph Theory, Numerical analysis and Integral Equations, Network Security and Cryptography, English to Hindi Translation System, Robotics and Level of Intelligence, Design and Implementation of a J2EEE Web Services using Soap-Over JMS Protocol, Wavelet Theory and Composition,, Some Problems in Symmetric and Asymmetric Cryptography, Nanotechnology, Education Management, Higher Education Management. He has published around 40 research papers in International Journals of repute and National/ International conference proceedings. 44 M.Phil. as well as 29 Ph.D degrees have been awarded so far under his dynamic supervision.

