

Predicting Stress Levels of Secondary School Students' using Machine Learning Approaches

Dr. A.V.R. Mayuri ¹, K. Asha Rani ², Dr. C. Sreedhar ³, Dr. Farooq Sunar Mahammad⁴

¹ Assistant Professor, G Pulla Reddy Engineering College (Autonomous), Kurnool

² Assistant Professor, G Pulla Reddy Engineering College (Autonomous), Kurnool

³ Associate Professor, G Pulla Reddy Engineering College (Autonomous), Kurnool

⁴ Assistant Professor, Santhiram Engineering College, Nandyal

Article Info

Volume 83

Page Number: 8977 - 8983

Publication Issue:

March - April 2020

Abstract

Stress is an emotional feeling that arises from any event or thought that has adverse effects on children since they feel frustrated, angry, tensed or nervous. Stress leads to severe problems such as less energy levels, unenthusiastic, poor attention, anxiety, depression and many other mental disorders. The level of stress observed in students studying in secondary school is increasing day by day which leads to poor academic performance. In recent days, lot of schools came in to existence and they are following different strategies in making their students to score high percentages in the board examinations. It is understandable that all children will not have the same capabilities and learning behaviour. Normally some students are able to overcome the stressful situations but it is not possible for few since they differ in terms of demographic factors, mental ability factors, study patterns, determination and psychological factors. So, it is necessary to predict the stress levels of the secondary school students and proper intervention mechanisms should be adopted at proper time for high risk students. This study aims to make the future generation a stress-free generation.

Article History

Article Received: 24 July 2019

Revised: 12 September 2019

Accepted: 15 February 2020

Publication: 09 April 2020

Keywords; stress, secondary school, success, examinations, prediction, machine learning.

I. INTRODUCTION

Many studies have shown that uneasiness and pressurized environment leads to different health issues and psychological problems. It is necessary for a person to balance the pressure constraints because stress leads to many health issues like obesity, mental depression, cardiac problems, diabetes, breathing problems, loss of sleep etc. Stress is defined as a kind of mental suffering which sometimes becomes life threatening. Many cases have been reported from different parts of the world that student suicides are increasing day by day because of stress. If we are able to predict the stress levels of students as earlier either in short or long term, then it will be helpful to recover the students from committing suicides or getting mentally disturbed. In this paper, we perform an analysis on

different stress factors that influence mental disturbances using classification algorithms by considering sensitivity, specificity and accuracy as performance metrics.

Scope and Exploration of Research questions:

The objective of this study aims in identifying different stress levels in academics perceived by secondary school students and adoption of suitable intervention strategies at proper time would help students in having great academic success. In this study, we examined the effects of different factors like individual, family, school and peer factors in an exploratory manner. Based on those factors, relationships related to the mental health of adolescents like parents – children relationship, faculty - student relationship, peer - peer

relationship were established. Highlighted terms used in the literature like mental stress, anxiety, academic success, prediction, detection, depression etc., were considered for exploring the research questions. As a result, “the prediction of mental stress in secondary school students” has been framed. To support the phrase, identification of metrics and features has been done. First, metrics has to be devised that best describes the student’s stress. Second, the set of features to be considered for better prediction of student stress levels and third, counter measures that has to be devised to reduce the stress levels and improve the student academic performance.

Research questions are framed in the context of metrics such as a) how stress can be defined in different terms? b) What suitable metrics can be used to describe student stress ? c) What considerable features should be taken in to account in predicting stress ? d) What are the possible predictive modelling techniques used in predicting stress ? e) How different combinations of features and methods can address different levels of stress in student? and what is the quality of work on predicting student stress?

II. REVIEW OF LITERATURE

Education is considered as the major pathway to social success for more than a millennium. In the field of education, perceived academic stress differs across gender, family background, and social and economic status [14]. It is reported that females have more stress or pressure than males since they worry about academic failures. Poor performance in academics is considered as one of the major factors in educational stress. Students with low grades or less Grade Point Average (GPA) often experience higher stress [11].

For about a century, many literature works [1][3][6] focussed on determining the stress levels in the learning environment and develop predictive models to identify the high-level stress in students. Many

scholarly literatures [7][9][13][14] exist on finding the factors that induce stress in students using various stress questionnaire.

A stream of research studies [2][10][12] aims in increasing students’ academic success by creating stress-less environment. The studies concentrated on finding out the difficulties faced by the students mentally. A proactive system was developed which gives inputs to the parents, guardians, instructors, mentors, counsellors for taking necessary actions to reduce stress in their wards.

Many research studies were about the severity of adolescent stress that has to be treated seriously and preventive interventions to be taken for mental health promotion. Many studies were focussed only on factors that influence stress such as pressurized study environment, heavy study load, poor relationships with parents, teachers and peers etc.

Few literatures [2][5][8][15] show that increase in stress leads to many adverse health problems like depression, suicide, heart failures etc. and the stress is measured in terms of heart rate, electromyography, electrocardiography, galvin skin response, electrodermal activity, blood oxygen saturation levels, respiration etc.

In paper [5], under four stressful conditions, stress was predicted using SMO, Bayesian Network, J48 algorithms. In the paper [4], it was found that stress was less during beginning of the semester and high during end of the semester since the students have to face the end examinations.

III. DATA PREPARATION

In this study, we considered a dataset of 186 secondary school students who were ready to face the board exams very nearer. The questionnaire includes 10 questions on demographic factors that includes gender, age, year of study, type of school, location of school, location of family, family level of income, parents’ occupation, parents’ education details, siblings studying in same school. Some of

these factors are considered as mediating factors because they indirectly influence stress on young minds. The Educational Stress Scale for Adolescents

(ESSA) questionnaire was used to estimate the stress level perceived in academics.

Factors	Description	Possible Values
Demographic factors	Gender	1-Male 2-Female
	Age	1- (7yrs-10Yrs) 2- (11Yrs to 14Yrs)
	Type of School	1-State Board 2- CBSE 3- ICSE
	Location of School	1- Rural 2-Urban 3- Sub Urban 4- Sub Rural
	Location of Family	1 - Rural 2 – Urban 3- Sub Urban 4- Sub Rural
	Family Level of Income	1 - Lower class 2 – Middle class 3 – Upper class
	Fathers Educational Background	1- Primary 2-Secondary 3- High School 4- Undergraduate 5- Post Graduate
	Mothers Educational Background	1- Primary 2-Secondary 3- High School 4- Undergraduate 5- Post Graduate
	Fathers Occupation	1- Agriculture 2-Others
	Mothers Occupation	1- Agriculture 2-Others
	Details of Siblings studying in same school	1- Yes 2- No
Personal Factors	Self-efficacy	1- Never 2- Sometimes 3- Always
	Health status	1- Poor 2-Moderate 3-Normal
	Body Mass Index (BMI)	1- Obese 2- Overweight 3- Normal 4- Underweight
	Physical Exercise	1-Yes 2-No
	Internet & Mobile Phone Usage	1- Yes 2 -No
	Playing Video Games	1 - Yes 2 – No

Table 1: Factors with descriptions and possible values

The Questionnaire contains 16 questions rated based on a 5-point Likert scale - 1 (Strongly disagree), 2 (disagree), 3 (neutral), 4 (agree) and 5 (Strongly agree). Based on the above total score is calculated. If the total is between 16 - 80 then it is considered as the high stress. The questionnaire mainly focussed on five factors. Study Pressure, Workload, Emotions (grades), Self-expectation and Dissatisfaction.

In addition to the above demographic and ESSA factors, other variables like self-efficacy, health status, Body Mass Index (BMI), physical exercise, Internet & mobile phone usage, playing video games was also taken in to consideration only for the past 30 days.

Factors	Description
F1: Study Pressure	Pressure in daily learning, parental pressure, competition with peers, concern about the future
F2: Workload	Burden related to Homework, school work and examinations
F3: Emotions (grades)	Stressful emotions due to academic grades dissatisfaction
F4: Self-expectation	Stressful feelings when self-expectations fail to be met
F5: Dissatisfaction	Lack of confidence, concentration in academic study

Table 2: ESSA Factors and its Description

ESSA FACTORS	Junior level (7 yrs to 9 yrs)		Senior level (10yrs to 14 yrs)	
	Average Score	Average Standard Deviation	Average Score	Average Standard Deviation
F1: Study Pressure				
Question1	3.41	1.4	3.46	1.43
Question2	3.39	1.37	3.44	1.39
Question3	3.53	1.37	3.59	1.39
Question4	3.62	1.41	3.69	1.43
F2: Workload				
Question5	3.71	1.3	3.79	1.3
Question6	3.34	1.4	3.38	1.43
Question7	3.38	1.43	3.42	1.46
F3: Emotions on grades				
Question8	3.49	1.36	3.54	1.38
Question9	3.65	1.35	3.71	1.38
Question10	3.63	1.34	3.69	1.35
F4: Self-Expectation				
Question11	3.41	1.34	3.5	1.38
Question12	3.5	1.3	3.56	1.31
Question13	3.77	1.34	3.85	1.35
F5: Dissatisfaction				
Question14	3.74	1.34	3.82	1.34
Question15	3.6	1.37	3.66	1.39
Question16	3.71	1.28	3.78	1.28

Table 3: Average Scores and Average Standard Deviation for ESSA Factors (N = 392: Junior Level - 206 & Senior Level-186)

Stress levels are analysed and the students were categorized in to three clusters - highly stressed, moderately stressed and normal. The performance of the model was further improved using 10-fold cross validation approach.

ESSA Factors	Average Score (Junior Level)	Average Standard Deviation (Junior Level)	Average Score (Senior Level)	Average Standard Deviation (Senior Level)
F1: Study Pressure	13.95	5.55	14.18	5.64
F2: Workload	10.43	4.13	10.59	4.19
F3: Emotions on grades	10.77	4.05	10.94	4.11
F4: Self-Expectation	10.68	3.98	10.91	4.04
F5: Dissatisfaction	11.05	3.99	11.26	4.01

Table 4: Total ESSA Scores

Proposed Approach:

Machine learning approaches were employed to predict students' stress levels and provide some proactive intervention strategies at proper time to

reduce the stress levels so that high academic success can be achieved.

Different machine learning classification algorithms like SVM, K-NN, Naïve Bayes Classifier and

Random Forest were employed on dataset. The comparisons of these algorithms were performed using three performance metrics – sensitivity, specificity and accuracy.

IV. IMPLEMENTATION

Statistical approaches, weighted average score and weighted standard deviation are used for item scores in ESSA. Using independent sample t-test the possible variances between primary and secondary level students were calculated. The correlation matrix was established using Pearson's Correlation Analysis between ESSA item scores and study related factors to know whether they are strongly correlated or not. All significant variables that are found to be useful were included in predictive models.

The performance parameters are evaluated and the predictive model uses 10-fold cross validation since the dataset was small. In 10-fold validation approach, the performance of the model is increased by increasing the dataset which generalizes the behaviour and increases the efficiency of the model. The implementation is done in Python language.

Sensitivity (True Positive Rate):

Sensitivity refers to the capability of the model to determine the stress level appropriately.

$$\text{Sensitivity} = \frac{\text{TruePositive}}{\text{TruePositive} + \text{TrueNegative}}$$

Specificity (True Negative Rate):

Specificity refers to the capability of the model to determine the stress level erroneously.

$$\text{Specificity} = \frac{\text{TrueNegative}}{\text{TrueNegative} + \text{FalsePositive}}$$

Accuracy:

Accuracy refers to the capability of the model to the stress levels accurately.

$$\text{Accuracy} = \frac{\text{TruePositive} + \text{TrueNegative}}{\text{TrueNegative} + \text{FalsePositive} + \text{TruePositive} + \text{FalseNegative}}$$

Machine Learning Algorithms:

K – Nearest Neighbour:

It is a simple, supervised classification algorithm. It classifies whether a student stress level falls under highly stressed, moderately stressed or normal. If required, there may be some sort of compromise between the neighbours.

Support Vector Machine:

Given a labelled training data, this outputs an optimal hyperplane which categorizes new data.

Random Forest:

It is an ensemble approach that works by constructing a group of decision trees at training time and outputs the class as a mean estimate of separate trees.

Naïve Bayes Classifier:

It is a simple probabilistic classifier that is based on Bayes theorem with strong independent conventions between the features of the dataset.

V. RESULTS AND DISCUSSIONS

Here we calculated the performance metrics (Sensitivity, Specificity and Accuracy) for four different machine learning algorithms (K – NN, SVM, Random Forest, Naïve Bayes).

It was found that SVM outperformed all the other algorithms giving an accuracy of 89.92%, sensitivity of 78% and specificity of 98.2%. Table 5 shows the comparison of different machine learning classifiers in percentages.

Classification Algorithm	Sensitivity	Specificity	Accuracy
K – NN Approach	72	46	59.55
SVM Classifier	78	98.2	89.92
Random Forest Approach	98	68.65	81.25
Naïve Bayes Classifier	67.38	78	72.92

Table 5: Comparison of different algorithms in percentages (%)

CONCLUSION

The stress level of secondary school students was estimated using the ESSA factors with some additional significant mediating factors. We can use the results of the prediction model to help the students to identify their risk of stress levels as early as possible, so that possible intervention strategies can be followed at proper time since the student can change the study pattern and they can easily overcome the stress. Machine learning Classification algorithms were employed on 186 secondary school student's datasets. The performance metrics were sensitivity, specificity, and accuracy. Here we employed 10-fold cross validation technique since the dataset was small. The findings were SVM outperformed well when compared with other algorithms. This is very cost effective since we used ESSA for analysing and it improves the mental health of future pillars of our country.

REFERENCES

- [1] Jung, Yuchae, and Yong Ik Yoon. "Multi-level assessment model for wellness service based on human mental stress level." *Multimedia Tools and Applications* 76.9 (2017): 11305-11317.
- [2] Norizam, Sulaiman. Determination and classification of human stress index using nonparametric analysis of EEG signals. Diss. Universiti Teknologi MARA, 2015.
- [3] Adnan, Nadia, et al. "University students stress level and brainwave balancing index: Comparison between early and end of study semester." *Research and Development (SCORED)*, 2012 IEEE Student Conference on. IEEE, 2012.
- [4] Subahni, Ahmad Rauf, Likun Xia, and AamirSaeed Malik. "Association of mental stress with video games." *Intelligent and Advanced Systems (ICIAS)*, 2012 4th International Conference on. Vol. 1. IEEE, 2012.
- [5] Khosrowabadi, Reza, Chai Quek, Kai Keng Ang, SauWai Tung, and Michel Heijnen. "A Brain-Computer Interface for classifying EEG correlates of chronic mental stress." In *IJCNN*, pp. 757-762. 2011.
- [6] Subhani, Ahmad Rauf, WajidMumtaz, Mohamed Naufal Bin Mohamed Saad, NidalKamel, and AamirSaeed Malik. "Machine learning framework for the detection of mental stress at multiple levels." *IEEE Access* 5 (2017): 13545-13556.
- [7] Xu, Q., Nwe, T.L., Guan, C..Cluster-based analysis for personalized stress evaluation using physiological signals. *IEEE Journal of biomedical and health informatics* 2015; 19(1):275–281.
- [8] Ang, R. P., &Huan, V. S. (2006). Academic Expectations Stress Inventory : Development, factor analysis, reliability, and validity. *Educational and Psychological Measurement*, 66, 522-539.
- [9] Bao, C. (2006). Policies for compulsory education disparity between urban and rural areas in China. *Frontiers of Education in China*, 1, 40-55.
- [10] Bhatia, S. K., & Bhatia, S. C. (2007). Childhood and adolescent depression. *American Family Physician*, 75(1), 73-80.
- [11] Bjorkman, S. M. (2007). Relationships among academic stress, social support, and internalizing and externalizing behavior in adolescence.(Unpublished Ph.D's thesis).Northern Illinois University, Illinois, United States.

- [12] Brown, S. L., Teufel, J. A., Birch, D. A., & Kancherla, V. (2006). Gender, age, and behavior differences in early adolescent worry. *Journal of School Health*, 76, 430-437.
- [13] Huan, V. S., See, Y. L., Ang, R. P., & Har, C. W. (2008). The impact of adolescent concerns on their academic stress. *Educational Review*, 60, 169-178
- [14] Jones, R. W., & Hattie, J. A. (1991). Academic Stress amongst Adolescents: An Examination by Ethnicity, Grade, and Sex. Paper presented at the Paper presented at the Annual 28 Conference of the New England Educational Research Organization (Portsmouth, NH, April 24-26, 1991)
- [15] Putwain, D. (2007). Researching academic stress and anxiety in students: some methodological considerations. *British Educational Research Journal*, 33(2), 207 - 219.