INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

RESEARCH ARTICLE

OPEN ACCESS

**Received:** 26-06-2023
**Accepted:** 22-08-2023
**Published:** 15-09-2023

**Citation:** Chopra A, Jain TK, Saxena N, Tiwari A (2023) Predictive Analytics of a Community Survey Data by Artificial Neural Network - A Subset of Machine Learning . Indian Journal of Science and Technology 16(34): 2778-2788. https://doi.org/10.17485/IJST/v16i34.1589

* **Corresponding author**.

amittiwari992@gmail.com

**Funding:** None

**Competing Interests:** None



Published By Indian Society for Education and Environment (iSee)

**ISSN**
Print: 0974-6846
Electronic: 0974-5645

# Predictive Analytics of a Community Survey Data by Artificial Neural Network - A Subset of Machine Learning

**Alka Chopra[1], T K Jain[2], Nishant Saxena[3], Amit Tiwari[4]***

**1** Ph.D Scholar, Department of Management, Suresh Gyan Vihar University, Jaipur, 302017, Rajasthan, India
**2** Professor, Department of Management, Suresh Gyan Vihar University, Jaipur, 302017, Rajasthan, India
**3** Assistant Professor, Department of Electrical Engineering, Suresh Gyan Vihar University, Jaipur, 302017, Rajasthan, India
**4** Assistant Professor, Department of Mechanical Engineering, Suresh Gyan Vihar University, Jaipur, 302017, Rajasthan, India

## Abstract

**Objective:** To evaluate the capacity of artificial neural network modeling in quantification of relative contribution of various factors towards happiness index of university faculty members and to adjudge the degree of agreement with the results of descriptive statistical analysis under hard societal situation. **Methods:** A relational-research is conducted by descriptive statistics and ANN modeling with 93 variables, grouped into 24 major variables. The primary data are obtained through surveying after random convenient sampling with self-administered questionnaire based on five-point Likert scale. The study is conducted on 350 faculty members; 273 duly filled in questionnaires are received. The data stemming from varying perceptions of teachers is highly nonlinear. ANN is chosen for its capability to capture high nonlinearity; a gold standard method of comparison with learning tools like multiple regression or logistic regression reveals it superiority; sample size driven predictive uncertainty makes machine learning unsuitable. **Findings:** ANN modeling shows that the independent variable 'salary' has 46% negative weightage on attainment of happiness whereas, the given working condition related factors record a 45-48% weightage. Descriptive statistics corroborate ANN result, showing that salary, job satisfaction and work environment cause dissatisfaction by recording poor happiness index ~ 65%. **Novelty:** The novelty lies in implementation of ANN modeling on community survey generated data for identification of significant affecting happiness of faculty members. Consideration of 93 influencing factors grouped into 24 input variable and to examine if ANN prediction corroborates statistical test of significance adds credence to the novelty of the present approach hitherto unreported.

**Keywords:** Happiness index; Artificial Neural Network; descriptive statistics; Training; student's ttest; Job satisfaction

## 1 Introduction

The dynamics of changes in science and technology have recently brought about a significant alteration in the sociological and psychological makeup of individuals, community and organizations. This has hit the conventional education system and the accompanying changes in teaching learning systems have affected the satisfaction levels of the university faculty members. It is observed that the structural and conceptual changes in higher education system have influenced the perceptions of university administration and the other stake holders; accordingly, the uncertainties in change management have exerted its effect on the mental satisfaction level or precisely saying, the happiness level of the university faculty members. The recurring mismatch between the rate of technology driven societal change and the adaptation rate of human community has stimulated intensive researches relating the impact of socio-economic changes on the happiness levels of the work force[(1)]. It has transpired that mapping the interrelation between happiness index (HI) of university faculty members and the potential factors that can affect the HI is an educated game of prediction. A happy person can take an organization forward by way of inducing the same feeling of happiness to others thereby transforming the system to a vibrant working place[(2)]. Happiness being the sensory experience of pleasure is responsive to a number external stimuli[(3)]. Commensurate with Hedonic happiness, it is needed to identify the factors that make university faculty members gain an experience of more pleasure at less pain under any situation of constraints[(4)]. Once identified, it becomes possible to take corrective measures for getting rid of unpleasant situation. Reports on identifying the determinants of university teachers' happiness under societal constraints are well documented in literature[(5)]. In recent time, machine learning (ML) technique is successfully employed in predictive analytics of reliable data; when the data set is created through community-based survey, human behavioral factors tend to produce an asymmetrical distribution of data. This skewed data set poses limitation in the act of drawing inference. To take care of the problem of establishing a non-linear relationship between multiple inputs and a single output, artificial neural network can be a good option. From the forgoing discussion it seems to be challenge to accurately predict the quantitative effect of a large number of happiness determinants on the satisfaction of the university faculty members. Therefore, the present investigation proposes to adopt a novel artificial learning techniques for handling 93 input variables grouped into 24 major ones against one output, happiness index. Statistical analysis is invoked to explore authentication of the output of predictive analysis.

## 2 Methodology

With the aim to identify the factors which tend to influence the happiness index of university faculty members under situation of severe constraints (Covid 19 like situation), real life data on certain identified factors viz. work environment during pandemic, job security, overall job satisfaction, salary and incentives, work responsibility, social endeavours, social support, assessment of health in educational institutions etc. are collected by extensive survey work. A total number of 24 items supposedly impacting the happiness of university faculty members, are taken up as the elements of survey. Each of the factors consists of several subitems (around 4 for each item). Based on the major variables and its constituent subitems, a questionnaire is prepared for the survey work to find out the happiness index of university faculty members [Table 1 of appendix]. The questionnaire is scaled similar to five-point Likert scale. A random convenient sampling method is taken resort to gather the data based on the responses of intendant faculty members. The questionnaire was served to more than 300 faculty members from different universities within the state of Rajasthan. The survey responses are provided in terms of linguistic variables; these are 'strongly agree', agree, neutral, 'disagree' and 'strongly disagree'. The hypotheses have been that a 'strongly agree' to subitems confers maximum happiness in respect of the concerned question. Linguistic variables are assigned stochastic token. Unlike conventional scaling of 1 to 5 sequentially from strongly disagree to strongly agree a more realistic and pragmatic scale is used for ANN modeling (Table 1); however, this scale has no different bearing on the calculated happiness index as compared to commonly used convention of one to five. It may be noted that the conventional scale, 1 to 5 (strongly disagree to strongly agree) has been used for statistical analysis.

Artificial neural network is known to be very effective tool to determine the nonlinear input-output system in a variety of areas[6,7]. So, in order to identify and for ranking the significant features influencing the outcome of response data (state of happiness under constraints), ANN modeling is invoked. This, upon validation by unused data, is used to identify the significant factors that impact the happiness of faculty members. Moreover, it is thought that for a better introspection into the pattern of input-output (happiness index of university teachers) relationship, artificial neural network, a subset of machine learning can be gainfully employed. The ANN model predicted input-output relationship are subjected to comparison by conventional statistical analysis.

**Table 1. Scale used to determine the contribution of a factor towards happiness of faculty member**

| Response | Stochastic value ($X_i$ ) | Impact on happiness Index (HI) |
|---|---|---|
| Strongly disagree | -2 | Most unhappy HI~0 |
| Disagree | -1 | Unhappy; HI~25% |
| Neutral | 0 | Neutral: HI~ 50% |
| Agree | 1 | Happy; HI~ 75% |
| Strongly agree | 2 | Most Happy; HI~100% |

In the present scenario the survey data were created with a view to map the happiness index of the university faculty members as functions of some presumed features described by sub-features. For example, salary could be considered as a feature within which, 'Financial equity between work, qualification and experience' have been used as sub-feature. The results of analysis have shown that, in the case of research activities, the satisfaction of faculty members has been moderate whereas, the teachers do not feel very happy with the assignment like acting as the Chairman/Convenor in organizing research related event. As regard to work environment the unhappiness emanates from non-availability of the resources to perform a given task. Psychological well-being as factor has revealed that negative emotions affected the happiness of mind quite considerably. However, the worst situation in respect of happiness can be witnessed for the cases of salary and job satisfaction. Statistical analysis corroborates the ANN model prediction between the actual response variable values and the predicted values.

The effect of different parameters on the overall happiness index of the university faculty members have been numerically calculated on the basis of received responses in linguistic variables, which are converted to numerical values as per information provided in Table 1. The HI scores are normalized by the use of the relation:

$$Normalized\ HI\ due\ a\ specific\ factor\ = \frac{(Xi - Xmin)}{(Xmax - Xmin)} \tag{1}$$

This assigns the normalized scores from 0 to 1 where 0 denotes 'strongly disagree' response and 1 is for the response 'strongly agree'. The other responses follow accordingly. The normalized values are used as the input variables and the happiness index as the output variables. The crisp dataset created from the responses are used to map the input-output relationship underlying the research question of what has been the happiness index of teaching community per se under constraints due to a pandemic, that imposes a number predicament affront the university faculty members. For this purpose, a neural network model is created.

## 3 Results and discussion

**Artificial neural network (ANN):**

In the present work we explore the efficacy of ANN as the learning tool for diagnosing the most critical factors affecting the happiness index of the university faculty members. The study is facilitated by the success stories of similar analysis meant to determine the happiness levels of stake holders in different academic scenario. As an example, analysis of the world happiness report and to classify the most critical variables affecting the life happiness score, a neural network classifier has been used to deduce that the per capita GDP affects the happiness index of individual quite appreciably[8]. On the contrary, the objective of the present investigation to assess the quantitative effects of happiness determining variables operating in university system. It is noted that with 7 happiness determinants as input variables, a neural network modeling can be employed to map the overall happiness of the students in a private university. The number of variables can be extended to 32 input variables by incorporating sub variables into each of the major variable. It has appeared that the support from the family and the demographics were the most dominant factors to fix students' happiness and that the prediction is more accurate than a typical regression analysis [9]. Successful forecasting of the academic performance of the university students is also possible with the aid of semi-supervised learning method [10]. There is a great similitude between students' happiness mapping activity with that for the university teachers.

Another encouraging work on the development of a suitable neural network model to accurately predict the probability that a student will pass an examination can be seen to have used ten characteristic attributes as the input variables for a three-layer ANN [11]. The said model has no need to infringe students' personal information, but has successfully demonstrated that the ANN model is superior to Naïve Bayes algorithm, Decision trees/random forests or Support vector machine (SVM)/support vector regression (SVR) in terms of ability to predict the academic performance of students at 83% accuracy level.

Since the factors influencing happiness index of university faculty members and the extent of happiness actually obtained in work place is nonlinearly related, it is considered wise to take resort to ANN modeling for the analysis of community survey data; this will make it possible to identify the most significant parameters affecting the happiness index of the faculty members. In these circumstances it is presumed that a constrained situation as in pandemic may further introduce perturbation in the response dataset related to the happiness of university faculty members. However, reports on ANN modeling of community-based survey data for identification of critical determinants of HI of university faculty members is not abundantly available in literature. Again, it is no denying that the quality of a university is, by and large determined by the performance of teachers, which, in turn improves the quality of outturns. Amidst tremendous competition among the centers of higher learning, the teachers are quite often given to handle a large number of time limited assignments. This tends to make teachers lose enthusiasm in teaching profession, thereby degrading the quality of the university. Thus, accurate diagnosis of the factors detracting teachers' interest in the job is essential. Concurrently, it is also required that due remedial measures are suggested by a good analysis of survey data. In fact, various learning are now-a-days contemplated to learn more about the real-life situation. The linguistic data has been generated by the community survey and it needs to be assigned with stochastic tokens. After conversion of linguistic variables into numeric variables, a new dataset is created which consists of 93 variables. As stated earlier, each major variable contains more than one constituent variable; excepting two items, all other items (major variable) contain 4 constituent variables; combination of so-called minor variables make up a single major variable that takes part in deciding the happiness index of university teacher; for example, the major variable like job satisfaction is consisted of four component responses to four questions under the broad parameter 'job satisfaction'. Thus, 93 variables are reduced to 24 major variables and these 24 variables are used as the input variables of the neural network. The neural network used for the present investigation has the structure 24i — [12] $_{h1}$ $1_{o,}$ the subscripts having the conventional meaning.

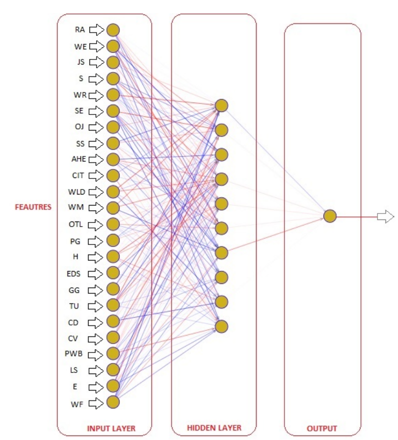


**Fig 1. Artificial neural network architecture used for the present investigation**

The neural network architecture used for the current study is presented in Figure 1. The 24 input nodes represent different factors which (in abbreviated form) are assumed (Table 2) to influence the happiness index of the university faculty members. As stated earlier, the normalized values of each variable are used. The feed forward back propagation learning algorithm, Lavenberg-Mrquardt (trainlm) available in ANN tool box of MATLAB software has been used for training. In the instant case Simulink tool is used for mapping of the inherent input-output relationship. It is known that each normalized input value is multiplied by a randomly chosen weight before being fed to the connecting nodes. In a hidden node all the incoming values (input * weight) coming from various input nodes are added and this new value is added with a bias value $[\sum_{i=1}^{i=n} wij\,xi + bi]$ for j=1 to 10. The value so derived is then operated with an approximator, known as transfer function; the corresponding numeric value is to enter the output node (a summing node); hence the values given by $[\sum_{i=1}^{i=n} wij\,xi + bi]_j$ for $j=1\ldots10$come out from each of the nodes in the hidden layer. These values are multiplied by randomly assigned layer weight and all these values are passed into output layer. The output node sums up all the incoming values and finally added with a bias value, $b_0$ to yield the output. Thus, the final output in the form of happiness index (HI) is to be given by HI= $\sum_{j=1}^{j=10} [\sum_{i=1}^{i=24} wij\,xi + bi] + b_{0\,\ldots}$…. (1). In the instant case we have used tanh transfer function and in the process of training mean square error is minimized [Figure 2]. As commonly practiced, 70% of the data has been used for training whereas 15% each are used for testing and validation. Figure 2 shows activity carried out in Simulink tool used for the present investigation. The general scheme of learning through neural network is shown in Figure 2 **(a)**.

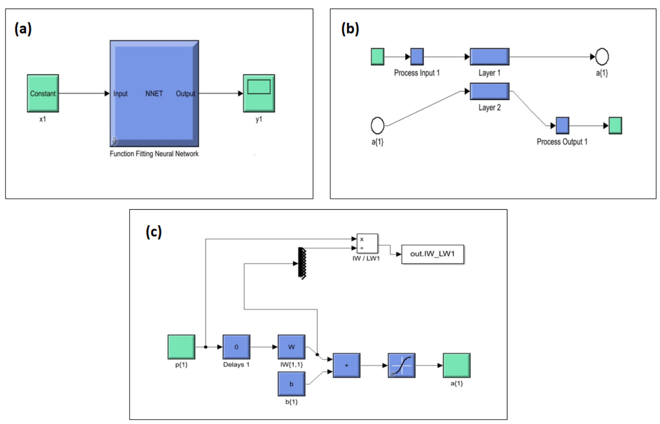


**Fig 2. (a) The scheme used by Simulink to carry out the training, testing and validation, (b) Representative functioning of artificial neural network, (c) The Simulink schematics for extraction of input weights**

The normalized values are used as input and Simulink uses the chosen algorithm and training parameters to give rise to the MSE values obtainable with progressing iteration.

Figure 2**(b)** demonstrates that each input node is connected to each of the hidden nodes. But a different weight is used for different hidden nodes; this means that one input variable is connected to 10 hidden nodes each with different weights. Upon following the theory of ANN modelling, it becomes clear that each of the input nodes share 10 weight values which ultimately follow the procedure of feed forward mechanism and finally determine the output. The major aim of the present work is to find out the relative weightage of various major variables in fixing the final output, that is the happiness index. This means it is intended to know how much a specific input variable contributes towards determination of happiness index. More precisely saying, we want to identify those variables which stand in the way of teachers' happiness. This entails the possible formulation of recommendation to maximize happiness index of university teachers and hence to improve the system efficiency. It is evident from Equation (1) that the final out is factually related to weights of individual variable by some nonlinear mathematical relation. Hence, the total contribution of an individual variable towards fixing the value of happiness index is somehow related to the total input weights available from an individual input parameter. Since the relational pattern between input weight and the output is identical for all the input variables, it is possible to derive the relative contribution of the input nodes; in other way, we can find out the relative dependency of output on the individual input from the total weight emanated from an input node.

Figure 2**(c)** shows the scheme of extraction of the weights from any input variable sent to ten hidden nodes. The results so obtained are suitably converted to dependency factor for each of the input variables. Each of the input node is assigned with 10 number of weights used for sending the weighted values to the 10 nodes in the hidden layer.

Let $S_0$ be the sum of all the weights from input variable 1 sent to various nodes and the corresponding value from input variable 2 is say, $S_1$.

The relative strength in contribution towards determination of output variable is to be expressed $S_0$ ~ $S_1$.

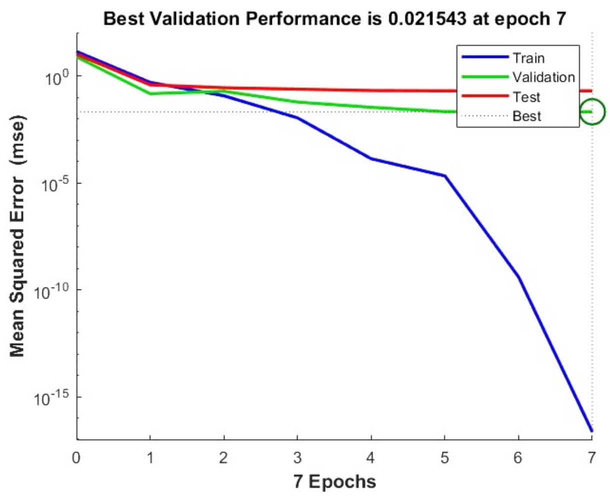


**Fig 3. Training, testing and validation curves of the implemented neural network**

The training curve in Figure 3 shows that a very low error ~$10^{-15}$ is achievable due to training the network by using Lavenberg-Mrquardt training algorithm. However, testing and validation does not follow the training curve throughout the process. This is not unusual in the cases where humane perception has been the key element towards data creation. Diverse demography, age and professional status do not allow 15% data, each fsor testing and validation, to maintain such consistency that these follow the training curve; in fact, this type of pattern recognition is not the motto of this investigation. It is observed that the best validation of 0.021543 is achievable after 7 Epoch. Incidentally the error level for such case is found to be around $10^{-2}$. This signifies that a generalized learning has been possible and it guarantees that the qualitative influence of an input variable on the network output does not alter even with diversified human factors.

Based on the architecture of NN used in the current investigation and following that $= \sum_{j=1}^{j=n} w_j x_j + b0$, where $x_j$ denotes the value of the output of jth hidden node. The weighted values of all these outputs from hidden nodes are passed into output node which sums up all these values and finally gets added with a bias value $b_0$. Taking this into consideration it is possible to save the NN output in the form of a final dependency equation in terms of $x_j$, 'j' ranges from 1-10. The final output equation for Neural Network comes out to be

$$y = -0.4186x_1 - 0.5446x_2 + 0.4405x_3 - 0.6972x_4 + 0.6623x_5 + 0.4039x_6 - 0.0027x_7 + 0.0856x_8 - 0.6272x_9 - 0.4062x_{10} + 0.15504$$

It is understood that all the $x_j$values (j=1 to10) are essentially determined by the weighted input values from all the 24 input variables. If the total weight values from individual input variable connected to all the 10 hidden nodes is determined, it is possible to predict the proportionate contribution of the concerned input variable towards determination of the final output value, i. e HI. Upon carrying out such exercise, the relative contribution of various factors is found out on the basis of total weights emanating from a single input variable for various hidden nodes and these are plotted in the form of a bar chart as a token of providing a comparative picture in respect of nature and quantum of influence of different variables (Figure 4). The abscissa of Figure 4 shows the abbreviated form of input variables in upper row, that is the parameters influencing HI and the lower row provides the corresponding value of the summed total weight of the input variable. The ordinate shows the relative contribution of different variables to happiness index. It is clear from Figure 4 that some of the input parameters have influenced the happiness index positively whereas some input variables have negatively impacted the happiness index. Its implication is that the factors having impacted HI negatively are those factors which are responsible for creating the feeling of unhappiness among the university faculty members. For ease of understanding, the relative contribution of various factors in order of increasing magnitude is shown in Table 2.

The survey data has been obtained during Covid 19 pandemic; during this time, a worldwide uncertainty in every aspect of human well-being was prevailing. It appears prudent to generalize the interpretation of the results of this investigation by stating

**Table 2. The abbreviations of the happiness determining factors arranged in order of increasing input weight (depicted in Figure 4)**

| Factors influencing Happiness | Abbreviations |
|---|---|
| Psychological well-being | PWB |
| Contribution of information technology | CIT |
| Salary | S |
| Online teaching–learning–assessment experience | OTL |
| Education | E |
| Personal Growth | PG |
| Good Governance | GG |
| Living Standards | LS |
| Time Use | TU |
| Work life balance | WLB |
| Work motivation | WM |
| Work Flexibility | WF |
| Social Endeavors | SE |
| Ecological Diversity and Resilience | EDS |
| Work Responsibility | WR |
| Health | H |
| Job security | JS |
| Cultural Diversity | CD |
| Social support | SS |
| Research Activities | RA |
| Overall Job satisfaction | OJ |
| Working Environment | WE |
| Community Vitality | CV |
| Assessment of Health in Educational Institutions | AHE |

that the negatively impacting factors are the causes of unhappiness of the university faculty members when the universities are to pass through a constrained situation, say, a global pandemic. Hence, the results of this investigation indicates that unless due attention is paid towards the so identified factors, the capability of the university faculty may be underutilized for the cause of the university and hence the efficiency of the system is to go down.

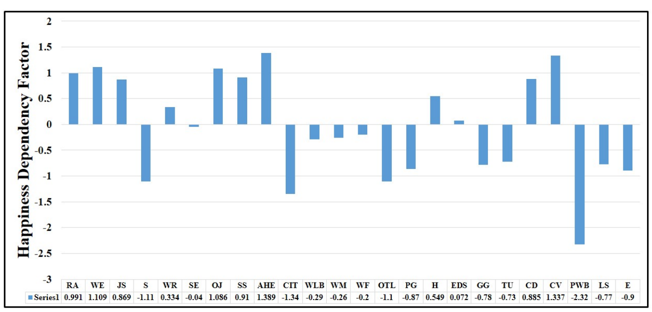


**Fig 4. The dependency of various factors on the happiness index**

It is apparent from Figure 4 that alarming factors adversely affecting happiness index are the salary, psychological well-being, contribution of IT infrastructure to conduct teaching, personal growth and health; of these factors, psychological well-being is one which is exclusive of individual's metal make up although a favorable make up may be aided by supports of different types from the university. Availability of infrastructure might have become important as data collection has been made just when the educational set-ups were getting them equipped for the purpose and this factor is not long-lasting as the problem was eradicated after sometime preparation time. So, salary and scope for personal growth are the two major factors which require to be duly attended by the employer specifically in private universities where financial aids from Government is not available. In general, the other prominent factors of interest viz. good governance and living standard are virtually related to the salary and job satisfaction components. It is apparent from Figure 4 that based on the maximum happiness value achievable due to a factor, ~ 1.38 (positive maximum), the weightage a factor affecting happiness index is described by maximum value- the negative weight value (~-1.10 for salary variable)/maximum value. This shows a negative weightage of salary component as a factor is more than 46%. The other concern is the infrastructural part enabling the faculty to obtain good working condition whose negative count lies within 46-49%. Thus, ANN study establishes that the major cause of unhappiness of university faculty members is the salary component and the working condition offered to the faculty members during the period of external constraints. It is prudent to compare the performance of the present work with a similar type of work due to Sinniah and coworkers[12] who have addressed the most pertinent research questions — how do the various extrinsic factors affect the satisfaction of the university teachers and what role the intrinsic factors play in modulating the teachers' motivation for excelling in performance. Like ours, the referred study uses the data obtained through a questionnaire-based survey among 343 teachers of a private university at Malaysia and within the limitation in number of variables used in the study as well as the lack of generalization in learning, the use of SPSS neural network algorithm could enable the Sinniah and others to verify that financial reward, personal growth and the intrinsic factors like work environment, code of conduct and teachers' freedom in work were most critical in impacting teachers job satisfaction. This obviously verifies the superiority of the present approach.

Moreover, the results of statistical analysis are summarized in Table 3.

**Table 3. Results of statistical tests of significance**

| Factors | Mean | SD | SE | t-score | Remarks |
|---|---|---|---|---|---|
| **RA** | 14.661538 | 3.84018 | 0.47644913 | 3.487335562 | $t_{cal}>t(table)$: H0 accepted |
| **WE** | 15.846154 | 3.8439994 | 0.476923 | 5.96774319 | $t_{cal}>t(table)$: H0 accepted |
| **JS** | 14.707692 | 3.5592699 | 0.44159676 | 3.867084516 | $t_{cal}>t(table)$: H0 accepted |
| **S** | 13.6 | 4.4299679 | 0.54962381 | **1.091655766** | $t_{cal}<t(table)$: **H0 rejected** |
| **WR** | 16.538462 | 3.9810201 | 0.49392309 | 7.163993902 | $t_{cal}>t(table)$: H0 accepted |
| **SE** | 15.692308 | 2.9084655 | 0.3608518 | 7.460979847 | $t_{cal}>t(table)$: H0 accepted |
| **OJ** | 13.492308 | 3.0690669 | 0.38077753 | 1.292901918 | $t_{cal}<t(table)$: **H0 rejected** |
| **SS** | 14.4 | 3.5897289 | 0.44537579 | 3.143412863 | $t_{cal}>t(table)$: H0 accepted |
| **AHE** | 13.523077 | 3.96177 | 0.49153474 | 1.064170969 | $t_{cal}<t(table)$: **H0 rejected** |
| **CIT** | 15.692308 | 3.8307383 | 0.4752777 | 5.66470502 | $t_{cal}>t(table)$: H0 accepted |
| **WLB** | 14.369231 | 3.7068304 | 0.45990452 | 2.97720712 | $t_{cal}>t(table)$: H0 accepted |
| **WM** | 14.153846 | 3.2020703 | 0.39727919 | 2.904370575 | $t_{cal}>t(table)$: H0 accepted |
| **WF** | 13.6 | 3.0117718 | 0.37366896 | 1.605699343 | $t_{cal}<t(table)$: **H0 rejected** |
| **OTL** | 16 | 3.4773996 | 0.43143916 | 6.953471784 | $t_{cal}>t(table)$: H0 accepted |
| **PG** | 14.969231 | 3.4237597 | 0.42478408 | 4.63583991 | $t_{cal}>t(table)$: H0 accepted |
| **H** | 14.276923 | 3.32592471 | 0.41264575 | 3.09447756 | $t_{cal}>t(table)$: H0 accepted |
| **EDS** | 13.815385 | 3.4769231 | 0.43138004 | 1.89017787 | $t_{cal}<t(table)$: **H0 rejected** |
| **GG** | 14.138462 | 3.5642538 | 0.44221511 | 2.574452953 | $t_{cal}>t(table)$: H0 accepted |
| **TU** | 7.4923077 | 1.746408 | 0.21667593 | 4.74 | $t_{cal}>t(table)$: H0 accepted |
| **CD** | 14.738462 | 3.2735627 | 0.40614922 | 4.280352938 | $t_{cal}>t(table)$: H0 accepted |
| **CV** | 15.169231 | 3.1892571 | 0.39568947 | 5.482155032 | $t_{cal}>t(table)$: H0 accepted |
| **PWB** | 14.323077 | 3.26336522 | 0.40488402 | 3.26779257 | $t_{cal}>t(table)$: H0 accepted |
| **LS** | 10.676923 | 2.5964929 | 0.32214552 | 2.852 | $t_{cal}>t(table)$: H0 accepted |
| **E** | 16.046154 | 3.4660142 | 0.43002658 | 7.083641273 | $t_{cal}>t(table)$: H0 accepted |
| **Total** | 341.92308 | 68.4780274 | 8.49603318 | 38.71490061 | $t_{cal}>t(table)$: H0 accepted |

**Descriptive statistics and tests of hypotheses:**

Table 3 presents the statistical analysis of the survey data. It shows the descriptive statistics of the individual variables which are supposed to bear the potential of influencing happiness index. The descriptive of the random sampling process that describes the standard error of the means is also shown in the same table (Table 3). Noting that an individual's responses encompass 93 answers and the maximum score for each answer can be 5, the happiness index of the studied population becomes 73%. This value is the cumulative effect of various major variables. It is important to understand if all the major variables contribute significantly towards the overall happiness index of university faculty members. It is apparent that the convenient random sampling made during the survey has yielded data from various persons and hence the data is independent of one another. It is presumed that random sampling has given rise to a normal distribution of data. The standard deviation data of each variable, delineate that there exists a good homogeneity in variance among the individual group of data (standard deviation of individual variable does not widely differ from one another). Thus, the survey data appears to be amenable to validation of hypothesis through student's t-test.

The problem deals with 24 individual variables contributing towards the response variable (output variable) in the form of happiness index of university teachers under situation of global constraints (say an ongoing pandemic). Therefore, it is required to know which of the factors have not been conducive towards attainment of happiness by a university faculty member.

Looking at the manner in which the stochastic values are assigned to the linguistic responses from the participants of the survey, it is pertinent to presume that at least a happiness index of 65% of maximum attainable value can be obtained from a particular variable to favor happiness index. For its clarification let us note that a neutral response is equal to a numerical value of 3 when it is assumed that strongly disagree has the credit of 1, whereas strongly agree has the credit of 5 (100%).

For a specific variable say research related activity (RA), there are four sub-variables; hence total maximum attainable score amounts to 20 when responses to all sub-variables are 'strongly agree'. Similarly, minimum score attainable by a variable is 4, for all the four responses being 'strongly disagree'. By the same logic the score is 12 if all four responses are 'neutral' and it is 16 if the four responses are all 'agree'. Thus, pragmatically a major variable does not cause unhappiness if it scores above 12. Hence, in the present work it is presumed that scoring 65% (~ 13) of the total maximum score (~20) is a token of positive contribution towards happiness index. This is also logically consistent with the calculated happiness index of 73%. That means if a variable response score exceeds 65% of attainable maximum, the concerned factor is not averse to happiness of the faculty member.

On the basis of above, the student's 't' test is carried out. For this purpose, the t-score of every major variable is calculated and are furnished in Table 1. The t-value for 64 degrees of freedom at $\alpha$ =0.05, (at 95 % confidence limit) '$\alpha$' having its usual meaning is found to be = 1.990.

The test of hypothesis for the factor 'salary' that exerts negative effect on happiness index is formulated as

H0: Happiness index of faculty members due to salary factor is at least 65%.

Thus, $H_0$: $\mu$ = 13 and $H_1$: $\mu \neq 13$.

The calculated value of $t_{n-1}$= 1.09, whereas the 't' value for 64 degrees of freedom with 95% confidence limit= 1.990 [$\alpha$=0.05 and n =65].

The calculated value is less than the actual t value and hence the hypothesis is rejected. This means that salary fails to positively influence happiness index and hence salary as a factor is a great concern for the happiness of university faculty members.

Thus, upon generalizing the results in Table 1, the hypothesis to be propounded is that happiness index due to each of the variables is at least 65% of the total that is 13 (except TU and LS).

Hence, our hypothesis is $H_{i0}$: $\mu$ = 13 [i=1-24; RA, WE. . . etc]

Else, $H_{i1}$: $\mu \neq 13$ [for $\alpha$=0.05 and n=65].

If we take the case of research related activity, i.e. RA, we write, $H_0$: $\mu$ = 13 and $H_1$: $\mu \neq 13$

The 't' value calculation results in $t_{n-1}$ = 3.487 against the value from student's t test table = 1.990.

Hence, the hypothesis is accepted; this means that happiness index due to RA is more than 65% and this factor does not come in the way of happiness of the teachers under global constraints.

It is found that calculated 't' value for the variables like salary (S), Overall-Job satisfaction (OJ), Assessment of Health (AHE), Work Flexibility (WF) and Ecological diversity (EDS) are less than the 't' values obtainable from the t-table at 95% confidence. This implies that the hypothesis is to be rejected for these variables. This means that the happiness index for these variables is less than 65% with default implication that these five factors do not aid in the attainment of happiness. For all the other variables, the 't' score is higher than 1.990 and hence the hypothesis is accepted. This also means that these factors give rise to higher than 65 % happiness index and thus do not adversely affect the happiness index and the hence the personal satisfaction of the teachers. It is observed that salary being the common component of dissatisfaction, the other predictive factors are mainly job satisfaction and its related factor. Health related issue has been identified as an important component, but we believe that it is time specific when everyone has been fearing the dreadful Covid 19 virus. So is the case with psychological well-being which under the pandemic threat has appeared as a matter of concern. Taking all these in consideration, it can be seen that the salary component and overall job satisfaction with its related factors are generic causes of dissatisfaction of university faculty members when there is an unwarranted socio-economic constraint. From the Table 3 it is possible to calculate the overall happiness index and it is found that happiness index of university faculty members is 73%. This value is quite low. It is reasonable to presume that unless the happiness index calculated on the basis of simple Likert's scale exceeds at least 80%, the prevailing mental state of the major stake holder of higher education that is faculty members is alarmingly poor. On the basis of this, a happiness score of ~65% or less is surely to be counted as a factor to cause unhappiness. On this basis, salary and factors related to job satisfaction counts the most unhappiness as the corresponding happiness scores lie between 65-68%. These fails in t-test too and this tends to support the above conjecture.

Moreover, on the basis of the understanding that there remains a gap in quantitative understanding about the relative effects, positive or negative, of the happiness influencing factors on the actual HI of teachers and that there is the need to conduct a study by taking a higher number of independent variables which actually have impact on the HI in real life scenario, the present study is supposedly rested on a more realistic foundation. Including Sinniah's ANN study there cannot be found any recorded attempt to evaluate the performance of ANN modeling by the use of a supplementary statistical tool so as to ensure the accuracy in prediction of relative strength of happiness impacting variable. As stated earlier, the present problem deals with community survey data where perception of different people determines the quality of data. There are reasons to believe that the collected data is of wider dispersion. The use of descriptive statistics to reinforce the sanctity of the predictive activity of ANN, is unique in character. Since by conducting t-test, comparison between any two HI influencing variables is possible, a better understanding about the influence of any factor on the HI of university faculty member has been gained. Since the statistical analysis provides the knowledge of the cumulative effect of all the factors in determining the overall happiness index, the use of such techniques adds to the novelty of the research. Thus, applying artificial learning technique on survey data to disclose the quantitative effect of the different happiness determinants for the university teachers and to assess the quality of the model prediction by employing statistical techniques is a novel approach hitherto unknown in the related area of research. Moreover, it is to be noted that unlike previously reported modeling of university teachers' happiness index, the use of highly efficient learning algorithm to map the relation between 93 input variable (influencing happiness index of university teachers) and a single output viz. happiness index is new by nature. The uniqueness of the present study lies in fixing the relative weightages of different variables while impacting teachers' happiness. This novel contemplation is essentially an act of prescriptive analytics which accurately identifies the factors needing employer's attention towards improving the satisfaction level of teachers and hence the university's performance credibility. Weighed against the previous research reports, the present work is the first of its kind as it invokes a quantitative approach to determine the relative magnitude of factors affecting the happiness index of university teachers.

## 4 Conclusions

The authors wish to conclude that the proposed neural network modeling can be used for prescriptive analytics and that it emulates the suggestive actions the university authority needs to contemplate for enhancement of its performance by way of eradicating unhappiness causing parameters. It is found that the suggested model can effectively analyze the community survey data comprised of 93 happiness determinants grouped into 24 major variables. Use of all the thinkable happiness determinants is the first of its kind and ensures its novelty; moreover, the suggested model can overcome the problem of generalization as experienced in a previous work[6]. The implemented ANN model deduces that salary related factors and the factors related to work place ambiance including infrastructural support are the major causes of unhappiness within the private university faculty members when the community is put to social constraints. This observation corroborates the result of previous work's finding that that financial reward, personal growth and the intrinsic factors like work environment, code of conduct and teachers' freedom in work are the most critical of all other happiness determinants[6].

The analysis of data used for the present study reveals that the negative weightage of salary component is higher than 46%. The other concern is the infrastructural part enabling the faculty to obtain good working condition and its negative count lies within 46-49%.

Descriptive statistical analysis shows that the overall happiness index of university faculty members under study is just 73%, a value, which is much lower than the normally expected value (~ 85% or higher). On basis of above, the calculated happiness scores of salary related variables and of the factors related to job satisfaction are found to lie within 65-68% or less. This signifies that these factors have been the major cause of teachers' dissatisfaction. This novel approach of authentic quantification (by statistical modelling supported ANN modeling) of the happiness impacting factors agrees with the qualitative observation in experimentation with teachers of a private university at Malaysia. This entices the authors to conclude that present study unveils the major dissatisfaction causing factors for teachers working in private universities at different parts of the world. The novelty of the work lies in a number of aspects; first there is no reported solution in weighing the quantitative effect of such a large number ~ 93 happiness determinants. Secondly, no ANN modeling of teachers' happiness index has used such a powerful learning algorithm as Lavenberg-Mrquardt. Moreover, evaluation of ANN performance by statistical modeling is conceptually new in the concerned field of research and hence it does add further to the novelty of the present work. It is also inferred that the reported study is not limited to mapping university teachers' happiness but also is capable to map the output variable against a large number of input variables obtained through community survey. Moreover, the study is useful to any employer dealing with employees of various categories and number; this is because a quantitative knowledge of factors causing dissatisfaction within employees, enables the employer to take such measures as to eradicate dissatisfaction among the employees, and hence, to secure enhanced performance level of the concerned organization. Thus, the study is generic in character and can reliably be employed for any societal problems needing serious attention.

## 5 Acknowledgement

The authors wish to acknowledge the support received from the authority of Suresh Gyan Vihar University while carrying out the study.

## References

1) Isa K, Jaes L, Damin ZA, Latif AA, Rahman AHA, Maslan N, et al. Determining Indicators Of Happiness Index Among University Staff. *International Journal of Engineering and Advanced Technology*. 2019;8(5c):726–730. Available from: https://doi.org/10.35940/ijeat.E1103.0585C19.
2) Arora RG. Happiness among higher education academicians: a demographic analysis. *Rajagiri Management Journal*. 2020;14(1):3–17. Available from: https://doi.org/10.1108/RAMJ-11-2019-0024.
3) Karakose T, Ozdemir TY, Papadakis S, Yirci R, Ozkayran SE, Polat H. Investigating the Relationships between COVID-19 Quality of Life, Loneliness, Happiness, and Internet Addiction among K-12 Teachers and School Administrators—A Structural Equation Modeling Approach. *International Journal of Environmental Research and Public Health*. 2022;19(3):1–20. Available from: https://doi.org/10.3390/ijerph19031052.
4) Martínez-Tur V, Estreder Y, Tomás I, Moreno F, Mañas-Rodríguez MA, Díaz-Fúnez PA. Not Every Day is Monday for Employees Confined due to COVID-19: Anticipatory Happiness Matters. *Journal of Work and Organizational Psychology*. 2022;38(3):201–211. Available from: https://doi.org/10.5093/jwop2022a12.
5) Casinillo LF, Casinillo EL, Aure MRKL. Economics of Happiness: A Social Study on Determinants of Well-Being among Employees in a State University. *Philippine Social Science Journal*. 2021;4(1):42–52. Available from: https://doi.org/10.52006/main.v4i1.316.
6) Sinha KK, Gupta MK, Banerjee MK, Meraj G, Singh SK, Kanga S, et al. Neural Network-Based Modeling of Water Quality in Jodhpur, India. *Hydrology*. 2022;9(5):1–21. Available from: https://doi.org/10.3390/hydrology9050092.
7) Dey S, Sultana N, Kaiser MS, Dey P, Datta S. Computational intelligence based design of age-hardenable aluminium alloys for different temperature regimes. *Materials & Design*. 2016;92:522–534. Available from: https://doi.org/10.1016/j.matdes.2015.12.076.
8) Khder MA, Sayfi MA, Fujo SW, Fujo. Analysis of World Happiness Report Dataset Using Machine Learning Approaches. *International Journal of Advances in Soft Computing and its Applications*. 2022;14(1):14–34. Available from: http://www.i-csrs.org/Volumes/ijasca/2022.1.2.pdf.
9) Egilmez G, Erdil NÖ, Arani OM, Vahid M. Application of artificial neural networks to assess student happiness. *International Journal of Applied Decision Sciences*. 2019;12(2):115–140. Available from: https://doi.org/10.1504/IJADS.2019.098674.
10) Widyaningsih Y, Fitriani N, Sarwinda D. A Semi-Supervised Learning Approach for Predicting Student's Performance: First-Year Students Case Study. In: 2019 12th International Conference on Information & Communication Technology and System (ICTS), 18-18 July 2019, Surabaya, Indonesia. IEEE. 2019;p. 291–295. Available from: https://ieeexplore.ieee.org/document/8850950.
11) Chavez H, Chavez-Arias B, Contreras-Rosas S, Alvarez-Rodríguez JM, Raymundo C. Artificial neural network model to predict student performance using nonpersonal information. *Frontiers in Education*. 2023;8:1–11. Available from: https://doi.org/10.3389/feduc.2023.1106679.
12) Sinniah S, Mamun AA, Salleh MFM, Makhbul ZKM, Hayat N. Modeling the Significance of Motivation on Job Satisfaction and Performance Among the Academicians: The Use of Hybrid Structural Equation Modeling-Artificial Neural Network Analysis. *Frontiers in Psychology*. 2022;13:1–14. Available from: https://doi.org/10.3389/fpsyg.2022.935822.