

Multi-Level Analysis System Education Data Based Meta Learning Classification

Sunita V. Lahane

PhD Student, Department of Computer Science,
Sant Gadge Baba Amravati University, Amravati, M.S. India

Dr. M. U. Kharat

Professor, Department of Computer Engineering,
MET Bhujbal Knowledge City,
Nashik, M.S. India

Abstract — Present paper is designed to validate the capacities of data mining techniques in context of higher education proposing a data-mining model for higher education system in the university. In this research work, classification technique is used to estimate student's performance. Since there are many techniques that are used for data classification, hence ANN method is used here. By this technique, ANN, which defines about the student's selection for college academic system. It provides prior information about the student selection and identification of college the dropouts who need special attention and allow the teacher to provide appropriate advising/counselling.

Keywords – Meta Learning; ANN; Seat Selection System; Classification.

I. INTRODUCTION

The introduction of information technology in several fields has lead the large number of data storage in different formats like records, files, documents, images, sound, videos, technical data and many new data formats. The data collected from different applications need proper method of extracting knowledge from large sources for better assessment making. Knowledge discovery in databases (KDD), often called data mining, aims at the discovery of useful information from large collections of data [1] [2].

The capability of to predict/classify a student's performance is very important in high dimensional educational environments. A very promising arena to achieve this objective is the use of Data Mining (DM) technique [3] [4]. In fact, one in every of the foremost useful DM tasks in e learning is classification. There are totally different instructional objectives for mistreatment classification, the duvet potential student teams with similar characteristics and reactions to a selected pedagogic strategy, to discover students misuse or game-playing, to cluster students UN agency are hint-driven or failure-driven and notice common misconceptions that students possess, to spot learners with low motivation and notice remedial actions to lower drop-out rates [5] [6]. To predict/classify students when using intelligent tutoring systems, etc. and there are different types of classification methods and artificial intelligent algorithms that have been applied to predict student outcome, marks or scores [7]. Calculating students grades from check scores using neural networks; predicting student educational success (classes that area unit booming or not) mistreatment discriminant operate analysis. Classifying students mistreatment genetic algorithms to predict their final grade; predicting a student's educational success using completely

different data processing methods; predicting a student's marks (pass and fail classes) using regression techniques [8] [9].

Educational data mining methods typically take issue from strategies from the broader data mining literature, in expressly exploiting the multiple levels of purposeful hierarchy in academic information. Strategies from the psychology literature unit usually integrated with strategies from the machine learning and data mining literature to realize this goal [11] [12] [17]. Education is a vital component for the betterment and progress of a rustic. It permits the individuals of a rustic civilized and well unnatural. Mining in academic atmosphere is named academic data processing, concern with developing new strategies to get information from academic information to investigate student's trends and behaviors towards education [13] [14] [18]. Lack of deep and enough information in higher academic system could prevents system management to realize quality objectives, data processing methodology will facilitate bridging this data gaps in education system [15].

In the field of data mining data about how students choose to use educational high dimensional software, it may be valuable an instantaneously consider data at the keystroke level, answer level, session level, student level, classroom level, and school level. Issues of time, sequence, and context also play significant roles in the study of educational high dimensional data [16]. This meta-analysis summarizes evidence from randomized educations matching the effect of educational involvements in which VPs were used either as another method or as preservative to usual programmer versus intercessions based on more methods that are out-of-date [19]. We accomplished also an assessment of the efficacy of VPs when used to accomplish results oriented to clinical reasoning or outcomes addressed to communication skills and ethical reasoning [20].

The paper is divided into different sections where Section II reviews some of the related work carried out on Meta learning classifier. Section III Describes developed meta learning architecture for Educational data selection system. Section IV gives implementation details. Section V shows results of meta learning process. Section VI concludes the paper and draws direction to future work.

II. RELATED WORK

Fabrizio Consorti *et al.* [21] have proposed a meta-analysis technique to performed the Effect Size (ES) from randomized studies comparing the effect of educational interventions in

which Virtual patients (VPs) were used either as an alternative method or additive to usual curriculum versus interventions based on more traditional methods. In this proposed methods, twelve randomized controlled studies were retrieved, assessing 25 different outcomes. Under a random-effect model, meta-analysis showed a clear positive pooled overall effect for VPs compared to other educational methods (Odds Ratio: 2.39; 95% C.I. 1.48-3.84). A positive effect has been documented both when VPs have been used as an additive resource (O.R.: 2.55; C.I. 1.36-4.79) and when they have been compared as an alternative to a more traditional method (O.R.: 2.19; 1.06-4.52). When grouped for type of outcome, the pooled ES for studies addressing communication skills and ethical reasoning was lower than for clinical reasoning outcome.

Lien Vanderkerken *et al.* [22] have proposed a technique for vocal challenging behavior (VCB) forms a common problem in individuals with autistic disorder. In this proposed methods, they evaluated the effectiveness of several psychosocial interventions applied to decrease VCB in individuals with autistic disorder, they conducted a meta-analysis of single-case experiments (SCEs). The overall treatment effect was large and statistically significant. However, the effect varied significantly over the included studies and participants. Examining this variance, evidence was found for a moderator effect of VCB type and intervention type, with, on average, the largest effects for interventions used to reduce VCB including stereotypical VCB and for interventions containing both antecedent and consequence components. Age, gender, primary treatment setting, publication year, and study quality did not significantly moderate the intervention effect.

Ramón Sagarna *et al.* [23] have proposed a fundamental question in the field of approximation algorithms, for a given problem instance, is the selection of the best (or a suitable) algorithm with regard to some performance criteria. They proposed multidimensional Bayesian network (mBN) classifiers as a relatively simple, yet well-principled, approach for helping to solve this problem. Precisely, they proposed an algorithm selection decision problem into the elucidation of the non-dominated subset of algorithms, which contains the best. This formulation could be used in different ways to elucidate the main problem, each of which could be tackled with an mBN classifier. They illustrated the feasibility of the approach for real-life scenarios with a case study in the context of Search Based Software Test Data Generation (SBSTDG). A set of five SBSTDG generators was considered and the aim was to assist a hypothetical test engineer in elucidating good generators to fulfil the branch testing of a given programmed.

Ningning Zhao *et al.* [24] have presented a study to explore the relationship between family socioeconomic status and mathematics performance on the base of a multi-level analysis involving a large sample of Chinese primary school students. A weak relationship was found between socioeconomic status and performance in the Chinese context. The relationship does not follow a linear, but a quadratic curve, implying that students from a disadvantaged family and higher socioeconomic background have a higher probability to attain higher mathematics scores. This could be explained because of Chinese cultural beliefs about education, exams and social class mobility. Moreover, the aggregated socioeconomic status at the school level seems to moderate in

the relation between individual SES and academic performance. This suggests that individuals from a disadvantaged family was achieved higher in the school with a higher family socioeconomic status than students who were enrolled in schools with a lower and average family socioeconomic status.

Alejandro Peña-Ayala *et al.* [25] have pursued a twofold goal, the first was to preserved and enhanced the chronicles of recent educational data mining (EDM) advances development; the second was to organized, analyzed, and discussed the content of the review based on the outcomes produced by a data mining (DM) approach. Thus, as result of the selection and analysis of 240 EDM works, an EDM work profile was compiled to describe 222 EDM approaches and 18 tools. A profile of the EDM works was organized as a raw database, which was transformed into an ad-hoc database suitable to be mined. As result of the execution of statistical and clustering processes, a set of educational functionalities was found, a realistic pattern of EDM approaches was discovered, and two patterns of value-instances to depict EDM approaches based on descriptive and predictive models were identified. The review concludes with a snapshot of the surveyed EDM works, and provides an analysis of the EDM strengths, weakness, opportunities, and threats, whose factors represent, in a sense, future work to be fulfilled.

Research Objective

This meta-analysis was designed to answer questions about the impact of instructional technology on postsecondary student's achievement and attitudes as both a combined collection of studies, and as two sub-collections of studies: 1) no technology in the control condition; and 2) some technology in the control condition. In addition, it looked at a set of moderator variables: 1) levels of education; 2) subject matter; 3) classroom/blended learning; 4) difference between treatment and control in technology use; and 5) pedagogical uses of technology.

III. PROBLEM DEFINITION AND CONTRIBUTION OF THE PAPER

A classical supervised classification problem consists of finding a function which, taking a set of random feature variables as arguments, predicts the value of a one-dimensional discrete random class variable. There exist scenarios, however, where more than one class variable may arise, so the extension of the classical problem to the multidimensional class variable case is increasingly earning the attention of the research community.

IV. EFFICIENT EDUCATIONAL DATA CLASSIFICATION THROUGH PROPOSED FEATURE EXTRACTION ALGORITHM WITH ANN CLASSIFIER

The ultimate target of this research is to design and develop a technique for email classification using naïve bayes classifier. The naïve bayes algorithm spam filtering is a probabilistic classification technique of email filtering which is based on bayes theorem with naïve independence assumptions. Let us consider each of the email can be illustrated by a set of features (attributes) $\{a_n\}$, where $1 < n > N$. Filtering of spam mails with naïve bayes by considering of all features is very difficult also, it need more time. In order solve this problem in this paper; we propose an efficient algorithm to select the significant features from the

available to filter the spam in efficient manner. The overall model of the proposed-mail spam classification system is

given in the following figure 1 and each part of the framework is elucidated concisely in the following sections.

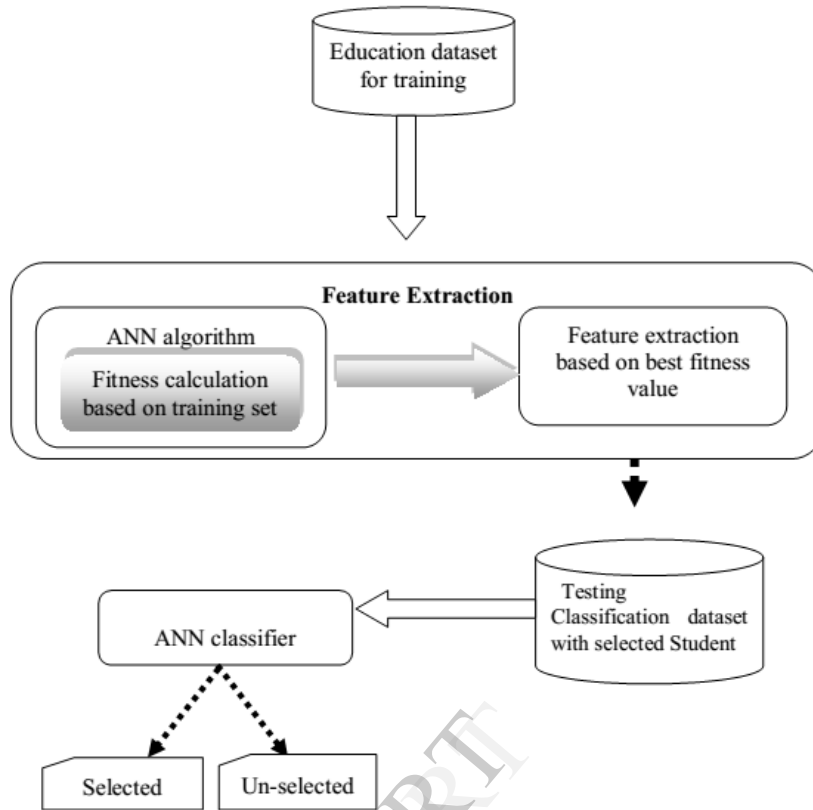


Fig. 1: overall model of the proposed-high dimensional classification system

A. Artificial Neural Network

In this paper, we use the artificial neural network for classification. Here, we training the artificial neural network

with various input and output layer weights and we update the weights of the neural network with the help of recently developed optimization algorithm [1]. The following figure represents the structure of artificial neural network.

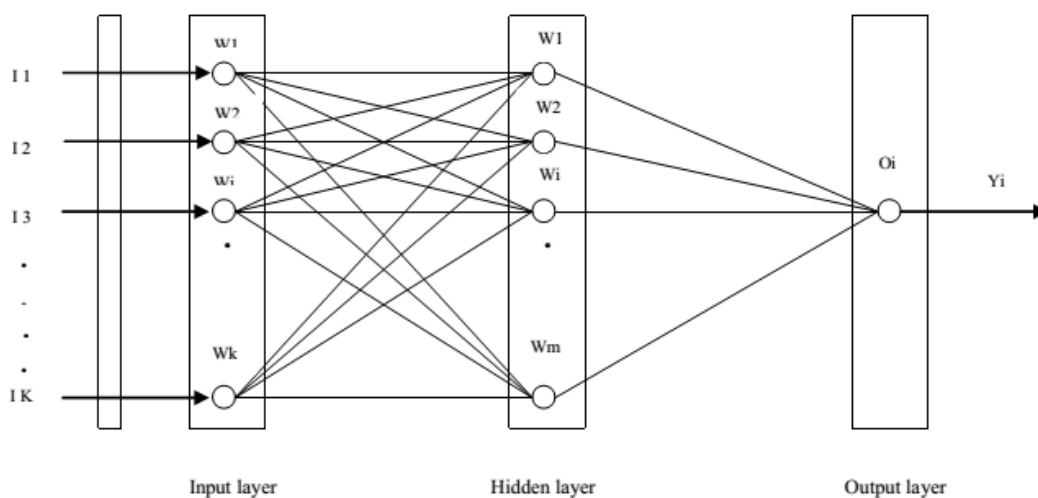


Fig. 2 shows that the general architecture of artificial neural network

Normally, an artificial neural network has an input layer, an output layer, with one or more hidden layers in between the input and output layer. ANN is an artificial intelligence technique that is used for generating training data set and

testing the applied input data. From the above figure $\{I_1, I_2, I_3, \dots, I_K\}$ are the input data which is pass through the input layer and output layer finally the ANN gives the

output Y_i from the output layer O_i . The set of inputs are given to the input layer, which consists set of weights $\{W_1, W_2, W_3..W_K\}$ and the result from the input layer B_i is given to the hidden layer, which also consists set of weights $\{W_1, W_2, W_3..W_M\}$ to calculate the output of artificial neural network.

B Assigning Weights to Input Layer and Hidden Layer

Normally, we initialize the input layer weight as randomly in the range of 0 to 1 and we calculate the output. In this paper, we use the ANN algorithm to choose the input layer and hidden layer weights. The ANN algorithm works based on the echolocation behaviors of ANNs. From the ANN algorithm, we get the values of many characteristics of ANNs such as velocity v_i , position p_i , frequency f_{\min} , f_{\max} wavelength λ_i , rate of pulse r_i and loudness A_0 from which we select the position p_i of the ANN as the weights of the input and the hidden layer of the artificial neural network. The dimension D (attributes) of the ANN's position p_i is decided based on the number of input layer K and number hidden layer M of the artificial neural network. The dimension of the each ANN's position (number of solution for each iteration) is calculated by the following eq.1.

$$D = K + M \quad (1)$$

Based on the number of training data of the artificial neural network, the algorithm finalizes the NN population $N(B_p)$. The NN algorithm generates $N(B_p)$ number of solutions at D dimensions randomly then the NN algorithm initializes the velocity v_i for each NN's position p_i , subsequently it also initialize the value of frequency f_{\min} , f_{\max} , λ_i , rate of pulse r_i and loudness A_0 . After the initialization of every parameter of the NN algorithm, the next step is to assign weights to input layer and hidden layer from the dimension D of the NN's position. The initial K number of attributes is assigned for input layer weights and next N number of attributes is assigned for hidden layer weights.

C. Training of Artificial Neural Network

Once the weights are assigned to input layer and hidden layer, the training of neural network gets start. The training of the artificial neural network has two main function the first one is, the input layer get the input and process it and gives output to hidden layer to process (i.e. hidden layer gets the input from the output of input layer). The second function done in the hidden layer, the hidden layer process the input from the input layer and gives the output to the output layer. Finally, the result of the ANN is get from the output layer.

1) *Output Calculation for Hidden Layers:* While training the artificial neural network, the input layer has two main sub functions to given the inputs for hidden layers, first one is the **basis function** and the second one is the **activation function** or transfer function which is given in the following equations (2), (3).

❖ The basis function

$$B^i = \sum_{k=1}^K (I_k \times w_k) \quad (2)$$

❖ Activation function

$$A^i = \frac{1}{1 + \exp(B_i)} \quad (3)$$

From the eq. (2) where B^i indicates basis function of i^{th} solution and the symbol I_k indicate the input and the symbol w_k represents the weight of the input layer. From the equation where (3) A^i indicates activation function of i^{th} solution.

2) *Output Calculation for Output Layers:* The ANN provides the input for hidden layer from the output of the input layer and the hidden layer process that data and donates the result to the output layer. The following equation (4) represents the calculation of hidden layer.

$$O_i = \sum_{h=1}^H (A^h \times w_h) \quad (4)$$

From the equation (4) where O_i indicates output function of i^{th} solution which is calculated in the hidden layer of the ANN. From the equation (4) A^h indicates the activation function output and w_h indicates the weight of the hidden layer.

Algorithm procedure

Input: Row student dataset with M number of attributes

Output: classified students seats

Parameters

$N(L_I) = K \rightarrow$ Number of input layers

$N(L_H) = M \rightarrow$ Number of hidden layers

$D \rightarrow$ Number of attributes of fitness solution

$P \rightarrow$ Number of NN population

$p_i^t \rightarrow$ Position of i^{th} NN at time interval "t"

$f_i^t \rightarrow$ Frequency of i^{th} NN at time interval "t"

$r_i^t \rightarrow$ Pulse rate of i^{th} NN at time interval "t"

$F_i \rightarrow$ Fitness of i^{th} NN

$p_i^* \rightarrow$ Best solution (NN position) at time interval "t"

$p_i^{t+T} \rightarrow$ Updated position of i^{th} ANN at time interval "t+T"

$w_k \rightarrow$ Weight of input layer

$w_h \rightarrow$ Weight of hidden layer

Begin

Step1. Get the number of input layers $N(L_I) = K$

Get the number of hidden layer $N(L_H) = M$

Calculate number of attributes (dimension) in fitness solution $D = K + M$

Get Number of P

Initialize ANN population p_i^t and velocity v_i^t

Define pulse frequency f_i^t at p_i^t

Step 2. Initialize pulse rate r_i^t and the loudness A_i^t

Call fitness F_i

Select best solution p_*

Generate new solutions p^{t+T} (equations 6, 7 and 8)

Step 3: Assign weights (solution) to input layers w_k and hidden layers w_h

Call fitness F_i

Arrange position of the solution based on fitness

Step 4. If $(r_i^{t+T} < \text{rand})$

Select a best solution p_{old} among them

Generate local best solution based on the best solution (equation 11)

Else

Generate new solution (equations 6, 7 and 8)

If $(A_i^{t+T} > \text{rand} \ \& \ F(p_i^{t+T}) < F(p_*)$

Accept that solutions as significant solutions $S(p)$

Else

Deny that solutions

Step 5: Calculate pulse increment ratio $\mu(sp_i)$ (equation 13)

Step 6: Calculate loudness decrement ratio $\omega(sp_i)$ (equation 14)

Increase pulse ratio $r(sp_i) = r_i^{t+T} + \mu(sp_i)$

Decrease loudness $A(sp_i) = A_i^{t+T} - \omega(sp_i)$

Arrange position of the significant solution based on fitness

Step 7: Select the best solution p_* from the significant solutions $S(p)$

Update loudness A_i^{t+T} and pulse ratio r_i^{t+T} (equations 9 and 10)

25. Go to step 12

Subroutine: Fitness

For each solution

1. Calculate basis function B^i (equation 2)

2. Calculate activation function A^i (equation 3)

3. Calculate output O_i (equation 4)

4. Calculate fitness F_i (equation 5)

End

3) *Calculation of Error (Fitness):* In this section, we evaluate the input data of the neural network and weight of the input layer and hidden layer based on the error value of the output. The error value must be less and it indicates how much the input data or weights of the hidden layer and input layer perfect. Here, error value is calculated based on the mean square error (MSE) value of the each iteration. The following equation helps to find the error (fitness) value of the solution.

Based on the error value of the solution only we update the solutions.

$$F_i = \frac{1}{N} \sum_{i=1}^N (\text{disired } O_i - \text{obtained } O_i)^2 \quad (5)$$

The above equation (5) which helps to find the error value of the solution or fitness value of the ANN's position. From the above equation (5) is the total number of training data of

the neural network data and *disired* O_i that indicates the original output of the given input data and *obtained* O_i that indicates the obtained neural network output. Likewise, for each training data ANN calculates the fitness value with the available solutions.

4) *Updation of Existing Solution*: The solutions are arranged in ascending order based on the fitness value of every input data of the ANN and the algorithm selects the first value ascending order as the best solution p_* of that iteration t . The solutions are generated in randomly in the NN algorithm since we update each solutions (positions) of the NN algorithm around the best solution since we found the best solution p_* initially subsequently we update the solution with the help of the following equations (6), (7) and (8).

$$p_i^{t+T} = p_i^t + v_i^{t+T} \quad (6)$$

$$v_i^{t+T} = v_i^t + (p_i^t - p_*) \quad (7)$$

$$f_i = f_{\min} + (f_{\max} - f_{\min})\beta \quad (8)$$

From the equation (6) p_i^{t+T} is represents the new solution, p_i^t represents the existing solution and v_i^{t+T} specifies that updated velocity. From the equation (7) v_i^t represents the existing velocity and p_* denotes best solution. From the equation (8) f_i specifies that frequency of the new solution, f_{\min} signifies minimum frequency, f_{\max} denotes that maximum frequency and β implies that random vector that belongs to 0 to 1. At the same time, algorithm updates the loudness A_i and pulse ratio r_i for every iteration with the help of following equations (9) and (10).

$$A_i^{t+1} = \alpha A_i^t \quad (9)$$

$$r_i^{t+1} = r_i^t [1 - \exp(-\gamma t)] \quad (10)$$

The updated new solutions are generated based on the above equations (9) and (10). The newly updated solutions are assigned as weights of the input layer and hidden layer of the artificial neural network and the ANN again calculates the fitness for each input values.

5) *Selection of Best Solution*: The newly generated solutions based on the value of pulse ratio r_i by generates the random value (*rand*). The solutions, which have the pulse ratio, less than random value $r_i < rand$ then such solutions are under group 1 and other solutions are under group 2. The solutions from the group 1 is updated with the help of following equation (11) and the other solutions (group 2) are updated based on the above equations (6), (7) and (8).

$$p_i^{t+T} = p_{old} + \varepsilon A^t \quad (11)$$

From the above equation (11) p_{old} is the best solution of the group 1 as well as the best solution of the group 1 is selected based on the fitness of the solutions among them. The value of ε is the random number which selected from the range of (-1 to 1). The value A^t is the average loudness of all

solutions in that iteration. The solutions from the group 1 and group 2 are combined into single group after the existing solutions are updated. Set of significant solutions $S(p)$ are selected from the combined group if the following condition (12) is satisfied moreover based on the fitness value of the solution we increase the pulse rate and reduce the loudness of the solution.

$$S(p) = A_i > rand \ \& \ F(p_i) < F(p_*) \quad (12)$$

From the above equation (12) $S(p)$ represents the significant solutions, which consists set of solutions $\{sp_1, sp_2, \dots, sp_n\}$. As per the best solution must has un qualified and qualified rate for that we change the selected and selected rate of the solutions for each solutions from the significant solutions $S(p)$ based on the following equations (13) and (14).

$$\text{Selected student ratio } \mu(sp_i) = \frac{F(sp_i) + r_i}{2} \quad (13)$$

$$\text{Total selected students ratio } \omega(sp_i) = \frac{F(sp_i) + A_i}{2} \quad (14)$$

After the calculation of selected student ratio $\mu(sp_i)$ and Total selected ratio $\omega(sp_i)$, we change the selected and Total selected of the significant solutions based on the following equation (15) and (16).

$$r(sp_i) = r_i + \mu(sp_i) \quad (15)$$

$$A(sp_i) = A_i - \omega(sp_i) \quad (16)$$

We arrange the significant solutions based on the fitness and we select and store the global best solution p_* and corresponds fitness after the modification of pulse ratio and loudness of those significant solutions. In further iteration, the value of global solution gets updates when the any solution has the better fitness value than the existing global solution.

V. RESULTS

A. Standardization and Descriptive Analysis

1. *Data Sets*: The selected high dimensional data utilized various measures of and different time intervals and session lengths. Hence, the obtained data were not immediately comparable. To solve this issue, the data were standardized. Using this method, we conducted a series of ordinary participant-specific regression analyses, whereby ANN was predicted by the condition. That way, the root mean squared errors were estimated. Subsequently, the raw data of each participant were divided by the participant's root mean squared error in order to get standardized data. Furthermore, before conducting the meta-analysis, we carried out a descriptive analysis to get more insight into the data. The obtained frequencies, means, standard deviations, ranges, and correlations of possible moderators and of descriptive variables are presented in Appendix A.

B. Selected Student Classification Results

From the whole selected set of data, some seats are taken for training and part of it are taken for testing purpose and then this procedure is repeated for the whole education

database. The results are evaluated with both the training algorithms on all the combination of sets, but due to space constraints, only some of the results are listed and compared. The specifications of selected students are given in table 1, 2,

3, and 4 with different colleges with two different branches. From the results, NN algorithm proves superior, as it is taking advantage of classification methodology and searching technique for classification

. TABLE 1: RESULTS WITH COLLAGE A WITH 500, 1000 AND 5000 STUDENT DATA.

	ANN Algorithm						
	Total	CS	NCS	CS	NCS	CS	NCS
MHSC	2	2	0	11	0	14	0
OTSC	2	0	2	0	5	0	5
NRISC	0	0	0	0	0	0	0
HASC	1	0	1	0	1	0	1
MHST	2	2	0	11	0	21	0
OTST	1	0	1	0	1	0	1
NRIST	0	0	0	0	0	0	0
HAST	0	0	0	0	0	0	0
MHBC	3	3	0	13	0	25	0
OTBC	2	0	2	0	2	0	2
NRIBC	1	0	1	0	2	0	2
HABC	0	0	0	0	0	0	0

TABLE 2: RESULTS WITH COLLAGE B WITH 500, 1000 AND 5000 STUDENT DATA.

	ANN Algorithm						
	Total	CS	NCS	CS	NCS	CS	NCS
MHSC	2	2	0	10	0	13	0
OTSC	1	0	1	0	4	0	4
NRISC	0	0	0	0	0	0	0
HASC	0	0	0	0	1	0	1
MHST	2	2	0	10	0	21	0
OTST	1	0	1	0	1	0	1
NRIST	0	0	0	0	0	0	0
HAST	0	0	0	0	0	0	0
MHBC	3	3	0	13	0	25	0
OTBC	1	0	1	0	1	0	1
NRIBC	1	0	1	0	2	0	2
HABC	0	0	0	0	0	0	0

TABLE 3: RESULTS WITH COLLAGE C WITH 500, 1000 AND 5000 STUDENT DATA.

	ANN Algorithm						
	Total	CS	NCS	CS	NCS	CS	NCS
MHSC	2	2	0	10	0	13	0
OTSC	1	0	1	0	4	0	4
NRISC	0	0	0	0	0	0	0
HASC	0	0	0	0	1	0	1
MHST	2	2	0	10	0	21	0
OTST	1	0	1	0	1	0	1
NRIST	0	0	0	0	0	0	0
HAST	0	0	0	0	0	0	0
MHBC	3	3	0	12	0	25	0
OTBC	1	0	1	0	1	0	1
NRIBC	1	0	1	0	2	0	2
HABC	0	0	0	0	0	0	0

TABLE 4: RESULTS WITH COLLAGE D WITH 500, 1000 AND 5000 STUDENT DATA.

	ANN Algorithm						
	Total	CS	NCS	CS	NCS	CS	NCS
MHSC	2	2	0	10	0	13	0
OTSC	1	0	1	0	4	0	4
NRISC	0	0	0	0	0	0	0
HASC	0	0	0	0	0	0	1
MHST	2	2	0	10	0	20	0
OTST	1	0	1	0	1	0	1
NRIST	0	0	0	0	0	0	0
HAST	0	0	0	0	0	0	0
MHBC	2	2	0	12	0	24	0
OTBC	1	0	1	0	1	0	1
NRIBC	1	0	1	0	1	0	1
HABC	0	0	0	0	0	0	0

VI. CONCLUSION

In this paper, we have presented an efficient technique to classify the student seats using ANN classifier. Initially, the input student data is given to the feature selection to select the suitable feature for shortlisted classification. The traditional ANN algorithm is taken and the optimized feature space is chosen with the best fitness. Once the best feature space is identified through ANN algorithm, the shortlisted student classification is done using the ANN classifier. The results for the shortlisted student detection are validated through evaluation metrics namely, sensitivity, specificity, accuracy and computation time. For comparative analysis, proposed spam classification is compared with the existing works such as particle swarm optimization and neural network for two datasets. Scalability is one of the features provided by the system where database is spread across the network and only decision tables (small in size) transferred through Meta learning agents are used to classify them.

As a future work, the system can be tested using number of different classification algorithms, So that their features can be combined and may prove useful for other applications student college selection. As features are being combined some of the demerits of different algorithms will also be combined, focus has to be given on this issue also.

ACKNOWLEDGEMENT

Authors thank Dr. V. M. Thakare, P.G. Department of Computer Science, Sant Gadge Baba Amravati University, Amravati, and Maharashtra; for his kind support in providing laboratory infrastructure facility required for the research work.

REFERENCES

- [1] Salvatore, Philip et al., "Meta learning agents for fraud and intrusion detection in Financial Information Systems.", In. paper Proceedings in International conference of Knowledge Discovery and Data mining, 1996.
- [2] S. Stolfo et al., "JAM: Java Agents for Metalearning over Distributed Databases," Proc. Third Int'l Conf. Knowledge Discovery and Data Mining, AAAI Press, Menlo Park, Calif., 1997, pp. 74-81.
- [3] Philip K. Chan, Wei Fan, Andreas L. Prodromidis, and Salvatore J. Stolfo, "Distributed Data Mining in Credit Card Fraud Detection", Proc. IEEE Int'l Conf. Intelligent Systems, Dec. 1999.
- [4] Jiawei Han, Micheline Kamber, "Data Mining Concepts and Techniques", pp. 279-328, 2001.
- [5] Salvatore J. Stolfo, David W. Fan, Wenke Lee and Andreas L. Prodromidis, "Credit Card Fraud Detection Using Meta-Learning: Issues and Initial Results", DARPA, 1999.
- [6] Todorovski L, Dzeroski S., "Combining Multiple Models with Meta Decision Trees.", In Proceedings of the Fourth European Conference on Principles of Data Mining and Knowledge Discovery. Springer, 2000 & Machine Learning Kluwer Publishers 2003
- [7] J.R. Quinlan, "Induction of Decision Trees, in Machine Learning", 106-181, 1986
- [8] Sam Mayes, Karl Tuyls et al., "Credit card fraud detection using Bayesian and Neural networks.", Proceedings in International Conference in KDD 2003.
- [9] J.R. Quinlan, "Induction of Decision Trees, in Machine Learning", 106-181, 1986
- [10] Zhang Yong, "Decision Tree's Pruning Algorithm Based on Deficient Data Sets", In Proceedings of the Sixth International Conference on Parallel and Distributed Computing, Applications and Technologies, 2005.
- [11] Arun Poojari, "Data Mining techniques", pp 150 -200, 1999.
- [12] Szappanos Tibor, Zolotova Iveta, "Distributed Data Mining and Data Warehouse", ASR '2005 Seminar, Instruments and Control, Ostrava, April 29, 2005.
- [13] William W. Cohen (1995), "Fast Effective Rule Induction, in Machine Learning": Proceeding of the Twelfth International Conference, Lake Tahoe, California, 1995
- [14] Sumit Goyal and Gyanendra Kumar Goyal, "Time-Delay Artificial Neural Network Computing Models for Predicting Shelf Life of Processed Cheese", Brain. Broad research in artificial intelligence and neuroscience, Vol. 3, No 1, 2012.
- [15] Sumit Goyal and Gyanendra Kumar Goyal "Application of artificial neural engineering and regression models for forecasting shelf life of instant coffee drink", International Journal of Computer Science Issues, Vol. 8, No 1, 2011.
- [16] Mailer, H. R., Dandy, G. C., "The use of artificial neural networks for the prediction of water quality parameters", Water Resources Research, Vol. 32, No.4, pp.1013-1022, 1996.
- [17] Lam, K.C. and Ng, S. T. and Hu, T. and Skitmore, M., "Decision support system for contractor pre-qualification: artificial neural network model", Journal of Engineering, Construction and Architectural Management, vol. 7, no. 3, pp. 251-266, 2009.
- [18] Garro, Beatriz A., Sossa, Humberto; Vázquez, Roberto A., "Artificial neural network synthesis by means of artificial bee colony (ABC) algorithm", In proceedings of the IEEE congress on evolutionary computation, pp. 331-338, 2011.
- [19] Cagdas Hakan Aladag, "A new architecture selection method based on tabu search for artificial neural networks", journal of expert systems with applications, vol. 38, no. 4, pp. 3287-3293, 2011.
- [20] Hatush, Z. & Skitmore, M., "Assessment and evaluation of contractor data against client goals using PERT approach", Construction Management and Economics, vol. 15, pp. 327-340, 1997.
- [21] Brijesh Kumar Baradwaj and Saurabh Pal, "Mining Educational Data to Analyze Students Performance", International Journal of Advanced Computer Science and Applications, Vol. 2, No. 6, pp. 63-69, 2011.
- [22] Saurabh Pal, "Mining Educational Data Using Classification to Decrease Dropout Rate of Students", International Journal of Multidisciplinary Sciences and Engineering, Vol. 3, No. 5, pp. 35-39, 2012.
- [23] Mohammed M. Abu Tair, Alaa M. El-Halees, "Mining Educational Data to Improve Students' Performance: A Case Study", International Journal of Information and Communication Technology Research, Vol. 2 No. 2, pp. 140-146, 2012
- [24] Manpreet Singh Bhullar and Amritpal Kaur "Use of Data Mining in Education Sector", In Proceedings of the World Congress on Engineering and Computer Science, San Francisco, USA, Vol. 1. pp. 6-9, October 24-26, 2012.
- [25] Bhise R.B., Thorat S.S., Supekar A.K "Importance of Data Mining in Higher Education System", Vol. 6, No. 6, pp. 18-21, 2013.



Aeronautical space research.

Sunita V. Lahane has received B.E. degree in Computer Engineering from Marathvada University, India in 1992, M.E. degree from Pune University in 2007. She is a registered Ph.D. student of Amravati University. She is currently working as Assistant Professor in Computer Engineering department in MIT, Pune. She has more than 10 years of teaching experience and successfully handles administrative work in MIT, Pune. Her research interest includes Data mining, Business Intelligence & Aeronautical space research.



Dr. Madan U. Kharat has received his B.E. from Amravati University, India in 1992, M.S. from Devi Ahilya University (Indore), India in 1995 and Ph.D. degree from Amravati University, India in 2006. He has experience of 18 years in academics. He has been working as a Principle of PLIT, Yelgaon, Budhana. His research interest includes Deductive Databases, Data Mining and Computer Networks.