# An Analysis of Student Behaviour for Attitude Prediction Using WEKA Tool

| **Sasikala. G** | **Dhanasekar.P** |
|:---:|:---:|
| Assistant Professor | Research Scholar |
| Department of Computer Science | Department of Computer Science |
| Adhiparasakthi College of Arts and Science, | Adhiparasakthi College of Arts and Science, |
| Kalavai, Tamil Nadu, India | Kalavai, Tamil Nadu, India |

*Abstract: A modern study is related to Education reveals that the civilization, attitude, aptitude, interest etc., totally differs from Government and Private school students.  On the basis of psychological studies, attitude is one most important factor of the Learning process. Our main aim is to show the comparison of the Attitude of Government and Private School Students towards training and development, School environment, studies, discipline, sports etc…   In this paper WEKA software is used for the classification of Government and Private School Students. Attitude, Discipline and Education is an essential part of our life.  The problem of present study entitled as "AN ANALYSIS OF STUDENT BEHAVIOUR FOR ATTITUDE PREDICTION USING WEKA TOOL". A comparative study of male and female students of the Government and Private schools significantly differs in the Attitude towards arts Education especially education discipline. The study reveals that there is a close relationship between the attitude and achievement in English learning as a personal experience of the researcher for last ten years, Government school students have positive attitude towards both education and discipline.  At the same time Private school students have better attitude towards education, sports, discipline etc.*

*Keywords: Government School, Private School, Students, WEKA tool, Data mining*

## I.   INTRODUCTION

We are in an age often referred to as the information age. In this information age, because we believe that information leads to power and success, and thanks to sophisticated technologies such as computers, satellites, etc., we have been collecting tremendous amounts of information. Initially, with the advent of computers and means for mass digital storage, we started collecting and storing all sorts of data, counting on the power of computers to help sort through this amalgam of information. Unfortunately, these massive collections of data stored on disparate structures very rapidly became overwhelming. This initial chaos has led to the creation of structured databases and database management systems (DBMS).

### 1.1 Attitude of Government And Private School Students

The Study compares the attitude of Government and Private school students towards Training & development, school environment, studies, discipline, sports etc. To compare the attitude of male and female totally differs from Government and Private school students. On the basis of psychological studies, attitude is one of the most important factors of the Learning Process.     The problem of the present study entitled as "An Analysis of Student Behaviour for Attitude Prediction Using WEKA Tool" A comparative study of Male and Female students of the Government and private  schools significantly differ in their attitude towards arts education especially education discipline. The study reveals that there is a close relationship between the attitude and achievement in English Learning. As a personal experience of the researcher for the last ten years, Government School students have positive attitude towards both education and discipline. At the same time private school students have better attitude towards education, sports and discipline etc.

### 1.2 Study objectives

- To compare the attitude of male and female students.
- To compare activities and Sports.
- To improve studying technology.
- Motivation, facilities etc.
- Dress code(Age/Gender)
- Infrastructure.

## II.   RELATED WORK

**A. Dinesh Kumar , Dr.V.Radhika Sri Krishna Arts and Science College, Coimbatore, India** [1] explained that Data mining is also called Knowledge Discovery in databases (KDD), is the field of determine useful information from

huge amounts of data. Data mining has been applied numerous arenas including bank, medical, census data, airline passenger records and super market, bioinformatics, business, genetics etc., Currently the data mining techniques has been used in educational environments. Applying data mining technique in Educational setting is called as educational data mining (EDM).Educational Data Mining is an emerging discipline, concerned with developing methods for exploring the unique types of data that come from educational settings, and using those methods to better understand students, and the settings Educational data mining (EDM) is concerned with developing, researching, and applying computerized methods to detect patterns in large collections of educational data that would otherwise be hard or impossible to analyze due to the enormous volume of data within which they exist.

**Richard A. Huebner Norwich University[2]** describes Oblinger (2007) defined academic analytics as the use of statistical techniques and data mining in ways that will help faculty and advisor s become more proactive in identifying at-risk students and responding accordingly.

**Smita M.Tech Student Amithy University, Priti Sharma Assistant Professor, Amity University [3]** describes Data mining means collecting relevant information from unstructured data. So it is able to help achieve specific objectives. The purpose of a data mining effort is normally either to create a descriptive model or a predictive model .

**Hardeep Kaur Master of Technology, Department of CSE, GNDU RC, Jalandhar, India [4]** explained Data mining is necessary process where various intelligent techniques are applied for extraction of useful patterns. Data mining comprises five major elements: Extract, convert, and load transaction data into data warehouse system. Storage and then management of this data in a multidimensional database system. Provide access of this data to information technology professionals and business analysts. Analysis of data using application software. Present the data in a useful form, such as a table or graph. Analysis and Visualization of Data It is used to highlight meaningful information and support decision making.

**Rajni Jindal and Malaya Dutta Borah Department of Computer Engineering, Delhi Technological University, N. Delhi, India [5]** describes The key components of EDM are Stakeholders of Education, DM Methods-Tools and Techniques, Educational data, Educational task and Outcomes which meet the Educational objectives **Stakeholders** Considering primary to higher education, major stakeholders of education can be divided in three groups: International Journal of Database Management Systems ( IJDMS ) **Primary group.** This group is directly involved with teaching and learning process. E.g.: Students (learners) and Faculties (teachers / learners, educators etc.) **Secondary group**. This group is indirectly involved in the growth of the institution. E.g.: Parents and Alumni. **Hybrid *group*.** This group is involved with administrative/decision making process e.g.:Employers, Administrator/Educational Planner, and Experts.

## III. PREDICTION

The current study involves three main phases prediction system: learning, experience and operation phases. In the learning phase, the knowledge based on literature reviews and interview with experts and students.

**Naive Bayes:** The probability model for a classifier is a conditional model.
$p(C|F_1, \ldots, F_n)$ over a dependent class variable $C$ with a small number of outcomes or classes, conditional on several feature variables $F_1$ through $F_n$. The problem is that if the number of features $n$ is large or when a feature can take on a large number of values, then basing such a model on probability tables is infeasible. We therefore reformulate the model to make it more tractable. Using Bayes' theorem, this can be written

$$p(C|F_1, \ldots, F_n) = \frac{p(C)\ p(F_1, \ldots, F_n|C)}{p(F_1, \ldots, F_n)}.$$

Using Bayesian Probability terminology, the above equation can be written as,

$$posterior = \frac{prior \times likelihood}{evidence}.$$

In practice, there is interest only in the numerator of that fraction, because the denominator does not depend on $C$ and the values of the features $F_i$ are given, so that the denominator is effectively constant. The numerator is equivalent to the joint probability model. $p(C, F_1, \ldots, F_n)$

**MLPA (Multi Layer Perception Algorithm)**

In feed-forward neural networks the neurons of the first layer forward their output to the neurons of the second layer, in a unidirectional fashion, which explains that the neurons are not received from the reverse direction.A kind of feedforward neural network mechanism is the Multi-layer Perceptron Neural Networks A MLPNN can be described as a feed-forward artificial neural network model that is Intelligent and Effective Diabetic retinopathy capable of mapping sets of input data onto a set of appropriate output. It is an alteration of the Prediction System Using Data Mining and Artificial Neural Network .once it weights them with the strengths of the respective connections j$_i$ w from the input layer and determines its output j y as a function f of the sum, given as    j ( j$_i$ i ) y = f ΣW X.At this instant it is possible for f to be a simple threshold function such as a sigmoid, or a Hyperbolic tangent function. The output of neurons in the output layer is determined in an identical fashion.

**Back-Propagation Training**

The back-propagation algorithm can be employed effectively to train neural networks; it is widely recognized for applications to layered feed-forward networks, or multi-layer perceptrons [46]. The BP algorithm is capable of adjusting

the network weights and biasing values to reduce the square sum of the difference between the given output (X ) and an output values computed by the net (X ') with the aid of gradient decent method as follows:

SSE = 1/ 2 NΣ (X – X')^2

Where N is the number of experimental data points utilized for the training.

**J48 Decision Trees**

A decision tree is a predictive machine-learning model that decides the target value (dependent variable) of a new sample based on various attribute values of the available data. The internal nodes of a decision tree denote the different attributes, the branches between the nodes tell us the possible values that these attributes can have in the observed samples, while the terminal nodes tell us the final value (classification) of the dependent variable.The attribute that is to be predicted is known as the dependent variable, since its value depends upon, or is decided by, the values of all the other attributes. The other attributes, which help in predicting the value of the dependent variable, are known as the independent variables in the dataset.

## IV. RESULTS AND DISCUSSION

The attributes that are corresponding to the student data are collected and predicted the behaviour using WEKA tool software. In this prediction many different types of algorithms were experimented, but Naive Bayes algorithm only giving the absolute results which is matching very closely to the actual results.

**Prediction Table:**
**Prediction Model in Naive Bayes:**

| | | |
|---|---|---|
| Correctly Classified Instances | 198 | 99 % |
| Incorrectly Classified Instances | 2 | 1 % |
| Kappa statistic | 0.9832 | |
| Mean absolute error | 0.0179 | |
| Root mean squared error | 0.0547 | |
| Relative absolute error | 5.9785 % | |
| Root relative squared error | 14.1787 % | |
| Total Number of Instances | 200 | |

**Prediction Model in Multi Layer Perception**:

| | | |
|---|---|---|
| Correctly Classified Instances | 190 | 95% |
| Incorrectly Classified Instances | 10 | 5 % |
| Kappa statistic | 0.9363 | |
| Mean absolute error | 0.0249 | |
| Root mean squared error | 0.0962 | |
| Relative absolute error | 10.8492% | |
| Root relative squared error | 28.6995% | |
| Total Number of Instances | 200 | |

**Prediction Model in J48 Perception:**

| | | |
|---|---|---|
| Correctly Classified Instances | 123 | 61.5 % |
| Incorrectly Classified Instances | 77 | 38.5 % |
| Kappa statistic | 0.5063 | |
| Mean absolute error | 0.189 | |
| Root mean squared error | 0.3074 | |
| Relative absolute error | 60.9976% | |
| Root relative squared error | 78.1322% | |
| Total Number of Instances | 200 | |

**Comparison of Algorithms**

Table 4.1 Comparison of algorithms

| | NAIVE BAYES | MULTI LAYER PERCEPTION | J48 |
|---|---|---|---|
| **ACCURACY** | 98% | 95% | 63.5% |

It is proven that navie bayes perception algorithm is most appropriate for predicting student's performance from the result. Navie bayes perception gives 99% of accuracy prediction which is relatively higher than other algorithms.
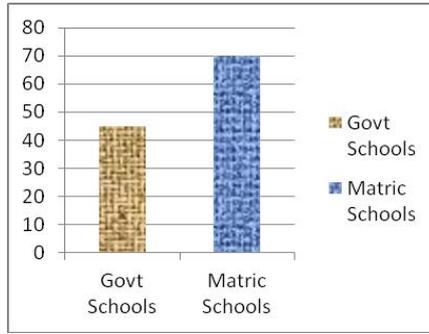
Figure 4.1 School classification

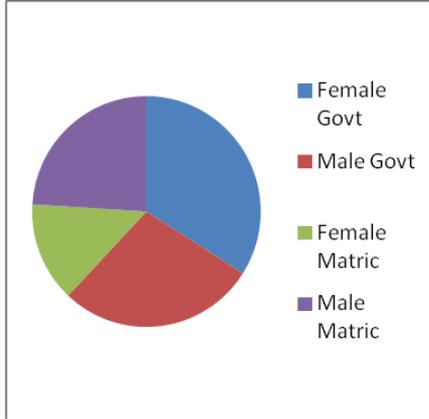Comparison of Male Students and Female Students in Govt & Matric School Students



Figure 4.2 Male & Female Students Comparison

**Analysis:**
Based on the gender analysis in Govt. & Matric Schools
Female Govt Schools : 34%.
Male Govt. Schools : 28%
Female Matric School : 14%
Male Govt. School : 24%
As per analysis Female students are greater than Govt Schools, Male students are greater than Matric Schools
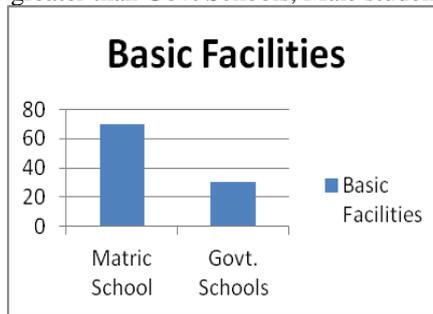


Figure 4.3 Basic Facilities of Govt. and Matric Schools

When comparing to Matric Schools basic facilities toilet and infrastructure is more higher than Govt. Schools
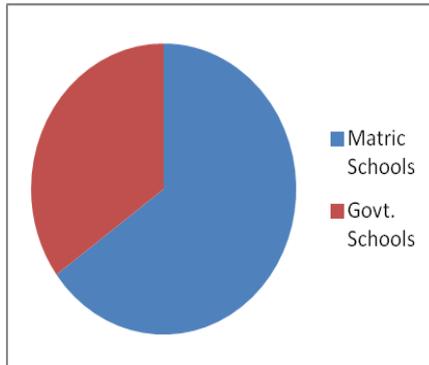


Figure 4.4 Overall Performance Comparisons

## V.  CONCLUSION AND FUTURE WORK

The system demonstrates the data mining based approaches that can be used to Analysis of Government and Matric School Students attitude

Data are collected from various types of schools through questionnaire. Finally, we have provided an efficient approach for the extraction of significant patterns from the collected data for the efficient prediction of Government and Matriculation School Students attitude.   The pre processed data was classified with Classification algorithm using Weka software. Thus using the Weka tool we performed the classification of data and the three main phases learning, experience and operation phases were carried out. Thus the system provides the way of comparative solutions given by various students. We have got the result of Student attitude which is having in Government or Matricualtion Schools

Our main aim is to show the comparison of the Attitude of Government and Private School Students towards training and development, School environment, studies, discipline, sports etc…   In this paper weka software is used for the classification of Government and Private School Students usage among various sector employees using Data Mining Technique' Naive Bayes. The performanace between Government and Private school students,  The private school and Private students are the best in Tiruvannamalai District.

## REFERENCES

[1]     A.Dinesh Kumar , Dr. V.Radhika Sri Krishna Arts and Science College Coimbatore, India A Survey on Predicting Student Performance, (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (5) , 2014, 6147-6149

[2]     Richard A. Huebner Norwich University A survey of educational data-mining research

[3]     Smita M.Tech Student Amithy University, Priti Sharma Assistant Professor, Amity University  Use of Data Mining in Various Field: A Survey Paper. IOSR Journal of computer (IOSR-JCE) e-ISSN,Vol.16,2278-0661.

[4]     Hardeep Kaur Master of Technology, Department of CSE, GNDU RC, Jalandhar, India A Review of Applications of Data Mining in the Field of Education, (IJRCCE) International Journal Advanced Research in computer and Communication Engineering , Vol. 4 (4) , April – 2015 ,2278-1021

[5]     Rajni Jindal and Malaya Dutta Borah Department of Computer Engineering, Delhi Technological University, N. Delhi, India a survey on educational data mining and research trends, (IJDBMS) International Journal Database Management System,         Vol. 5 (3) , June  – 2013 ,2278-1021

[6]     Nikitaben Shelke, Shriniwas Gadage Computer Engineering, Savitribai Phule Pune University, Maharashtra, India A Survey of Data Mining Approaches in Performance Analysis and Evaluation, (IJARCSSE) International Journal of Advance Research in Computer Science and Software Engineering, Vol. 5 (4) , 2015, 2277-128X

[7]     Shubhi Tomar, Garima Verma, Anuj David, Megha Thakur, Miti Judith Ogorchukwu, Purnima Parmar, Prafulla Shriyan, Richa Jaswal  Department of Public Health Manipal University, India)  Study on stress and coping strategies among private and government high school children in Udupi Taluk, Karnataka-A cross sectional study,(IOSR-JNHS)-IOSR Journal of Nursing and Health Science Vol. 3 (2) , 2015, 2320- 1959.

[8]     Poonam Asst. Professor, Kshatriya College of Education, Kathgarh Road. Indora (H.P.) A Comparative Study of Environmental Awareness among Government and Private Secondary School Students, International Journal of Educational Planning and Administration, Vol.2 (2) , 2014, 2249-3093.

[9]     K.Umamageswari and  S.Niraimathi NGM College , India  A Study on Student Data Analysis Using Data Mining Techniques,  International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 8, August 2013 ISSN: 2277 128X

[10]    Sumesh. P. M. Prof. (Dr.) Abdul Gafoor. P. K. Attitude of Government and Private School Students towards Geography Education in Calicut District of Kerala – A Comparative Study, Journal of Education and Human Development, Vol. 2 No. 2, December 2013