

## A MODERNISTIC APPROACH USING DATA MINING IN PREDICTING HUMANISTIC BEHAVIOUR

D.Radha<sup>1</sup>, R.Jayaparvathy<sup>2</sup>, Sandhya.T.R<sup>3</sup>, Vaishnavi.Y<sup>4</sup>

<sup>1</sup>M.O.P Vaishnav College for women (Autonomous), Chennai, India

<sup>2</sup>SSN College of Engineering, Chennai, India

<sup>3</sup>M.O.P Vaishnav College for women (Autonomous), Chennai, India

<sup>4</sup>M.O.P Vaishnav College for women (Autonomous), Chennai, India

### **ABSTRACT**

*Data mining refers to the knowledge discovery in data base process (KDD) which is a procedure from which useful information is drawn and summarized into a specific pattern. In other words it is the process of extracting and gaining useful knowledge data. "Better sound through research". The paper centralizes on the prediction of humanistic behaviour and hence integrates and manipulates various ways that a person can predict the behaviour. It utilizes decision tree, one of the data mining models to classify the causes. The purpose of this paper lies in finding out how people predict the behaviour of another person amongst different traits and hence concluding on to which they give importance. The tool used to analyse the data set is Rapid Miner. This paper helps on to conclude with what feature a person predicts behaviour and is useful in psychological concerns.*

**KEYWORDS:** KDD, Behaviour analysis, prediction, decision tree, traits

### **I. INTRODUCTION**

Behavior might depend upon any environmental circumstances. A basic knowledge on how to predict the behavior is sometimes necessary for an individual. For instance, in an interview a person's behavior is first observed. Thus behavior determines who we are. There can be various areas in which this can be observed. On such note, this model constructs the basic bridge for which an individual can predict behaviour.

Paper focusses on people on to improve the level of prediction on others behavior.

Data mining mines the meaningful information in order to make the data more clear and to derive to the conclusion. This helps people/students to understand the concept in an efficient way.

There are two basic types called as the predictive and the descriptive task. Predictive task is where the value of one attribute is determined with the help of the other attribute. Descriptive task is a process of drawing patterns and relationships amongst the data.

This paper uses classification model which comes under the predictive task. Classification is used to exactly predict the target class i.e. the resultant. It classifies data based on the data set values.

A decision tree model is a pictorial representation of data that is created based on the training set values. The nodes represent the attributes, branches representing the conditions that have been analysed and the leaf representing the result set.

Section I deals with the literature review related to this topic. Section II deals with the system model is presented referring to the main task of the paper. Section III is the performance analysis and results which describes the representation and analyzation of the performance using a data mining tool and how the results are interpreted. Section IV describes the conclusion derived from the interpreted results. Section V deals with the future work that can be extended further to predict the behavior using the colour analysis.

## II. RELATED WORK

Behavioral Data Mining is a research that is emerging to have a better understanding and to determine from the level of perception of different people. Choosing certain circumstances under which a person will/can predict the behaviour of another person and mining it using an effective data mining technique to conclude on which trait a person predicts the same is what will be seen in the further description.

**A source for consumer Behavioral Analysis:** This paper by Abhijit Raorane & R.V.Kulkarni, studied the analysis of the consumer behaviour when purchasing and was mined using effective data mining techniques that improves the conventional method by identifying the psychological condition of the purchaser when purchasing.

**Behavior-Analytic Research On Dementia In Older Adults:** In this paper, Maranda A Trahan, Alyssa B Fisher, Nicole L Hausman & Sung Woo Kahng segregated the analysis of behaviour based on Dementia (disorder of mental processes caused by brain) determining it only for adults specifically.

**Use of Applied Behavior Analysis with An Autistic Adolescent:** A case study based on the behavioral analysis and assessing the behaviour of an adolescent male who are diagnosed with autism (a mental condition in finding difficulty in communicating etc...)

**Looking to the Future: Will Behavior Analysis Survive and Prosper?:** Alan Poling, studied current Behavioral analysis is doing well. This research concludes on whether the behavioral analysis in the future will survive and prosper based on the applied behavioral analysis that improves the quality of life.

## III. SYSTEM MODEL

Paper concentrates on how a person is being evaluated with his/her behavioral qualities based on various categories. Mainly evaluates to which category is given the highest priority with the help of the Rapid Miner tool. There are many ways in analysing one's behaviour. Some of the categories in which the attributes or the fields are chosen is as follows: appearance, communication, thoughts, favourite colour, facial expression, habits, moods, reaction/reacting, response and others (if any).

These fields may seem similar but each of them differs in a slight variation as that of the others. A cognitive process is taken in order to mine the data to deduce the conclusion, which is done using the tool. In general, behaviour is measured by feelings. Feelings are nothing but, sense that is felt by us and lasts longer than the emotions. Emotions can otherwise be said as reaction, since emotions are the expressions or reactions that are expressed out by a person as an outcome of a feeling. Emotions changes to ideas.

Most of the approaches concentrate on Behavioral analysis - in the fields of determining the behavior in specific areas such as gender, adults and adolescents amid data mining, but not in predicting how a behavior can be analyzed. This paper focuses on that idea.

The classification model is used to segregate the target label based on certain conditions. The data gained is collected from different people. Classification is otherwise coined as supervised learning, where a new data is classified based on the training data set.

In this attempt, the collected data is organized as a table consisting of rows and columns which holds the attributes and its corresponding data. The attributes are used in the form of abbreviations and is listed here:

APP	Appearance
COM	Communication
FCLR	Favourite colour
FEXP	Facial Expression
HBT	Habits
MD	Moods
RCT	Reaction
RES	Response
THGT	Thoughts
OTH	Others

**Fig: 1:** Attributes used

Each attribute was selected based on how people generally analyze behavior. This predictive analysis is a mere attempt to find a better solution, where people often might go wrong when it comes to this situation. Each attribute has its own value for studying other's behavior. Conversations also play a vital role in this. But conversations are different from communication.

A set of procedure generally termed as algorithm that is used to implement classification technique is called as classifier. There are two types of variables called the independent and the dependent variables. The value which can be determined based on the values of other variables is called the dependent variable. The dependents variables are the predicted values which is the outcome. While an independent variable is referred to as the explanatory variable.

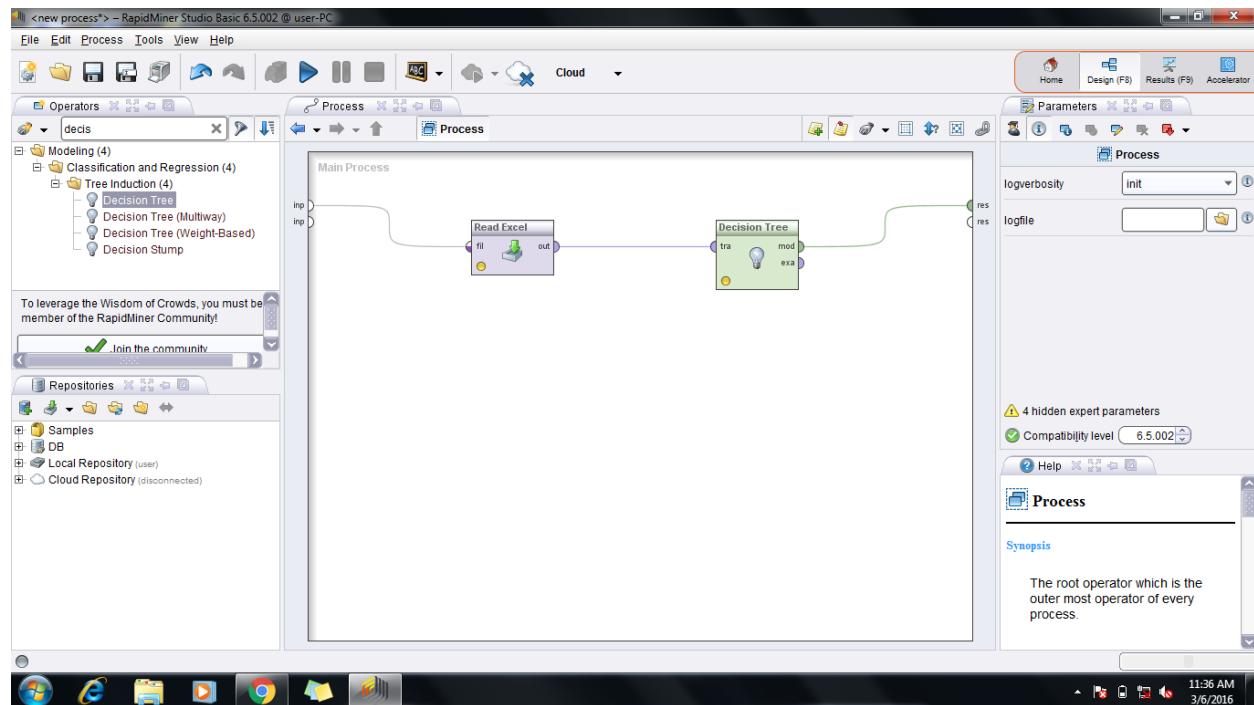
People's perspective differs from each other and the level of prediction differs too. This opinion from different kinds of people helps in determining how many people predict it highly. The target variable is the label attribute which predicts whether a person's prediction on behavior is up to the level and to which kind of attribute high priority is given. The independent variable is a text attribute with values as 'high' or 'low'.

If the value is high for a specific record, it signifies that their capability of prediction is high and similarly if the predicted value is low, it signifies that their capability in analyzing behavior is low.

The attributes here are rated on a scale basis and a survey is taken from a large set. While each attribute is taken on the whole of weight 5, people scale it accordingly from their level of perception. The total weight is estimated for each record. Also, total weight is calculated for each attribute and those attributes with the maximum value are taken into consideration. A sub average is then calculated for those attributes, as in the way of giving priority for those categories where by specifying some conditions to predict the label attribute as 'high' or 'low'.

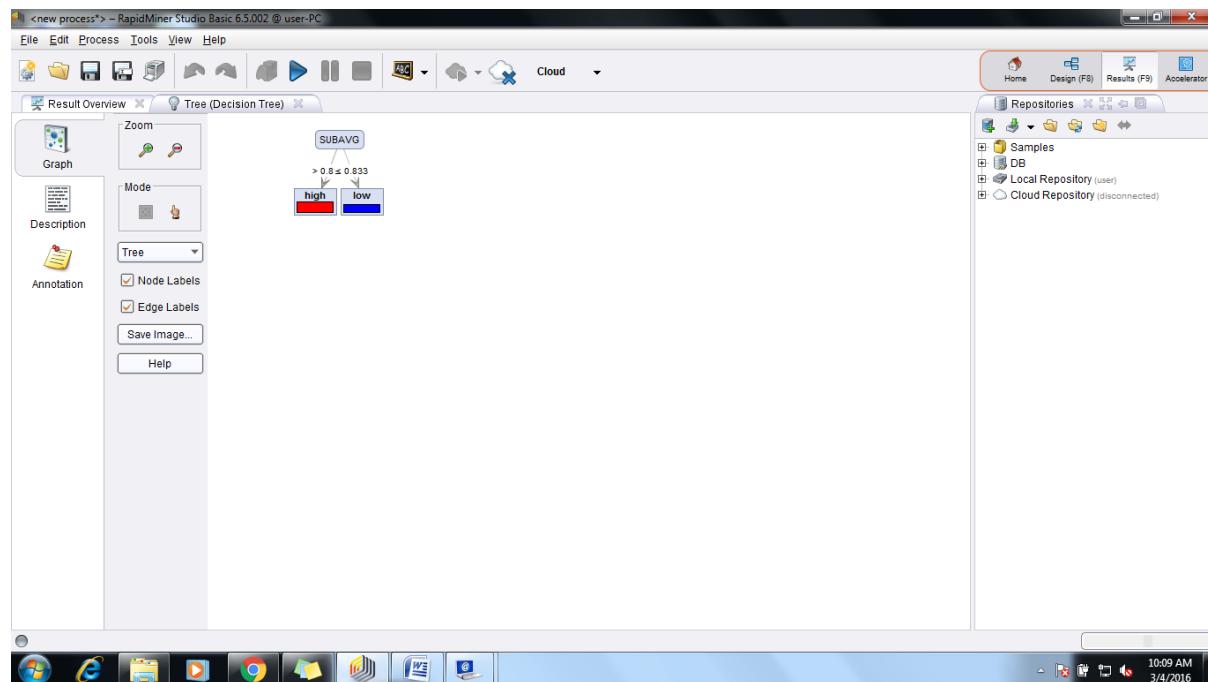
A constraint is set where a threshold value is fixed for the model. Those records whose sub average value is greater than the threshold value signifies that the prediction level is high, thereby concluding that people focus on those categories in which the weight is above the threshold value.

#### IV. PERFORMANCE ANALYSIS AND RESULTS



**Table:1**

### Decision tree:



**Table: 2**

The decision tree is generated by a data mining tool called Rapid Miner. The tree built is based on the given data set values and the main attribute is determined as the root of the tree. Depending on the values, it determines the prediction level based on two categories.

Data set:

Row No.	LABEL	APP	COM	FCLR	FEXP	HBT	MD	RCT	RES	THGT	OTH	TOTAL
1	low	3	5	1	4	2	3	3	5	4	1	31
2	high	4	3	2	2	5	2	4	5	3	2	32
3	low	4	3	1	5	3	2	3	5	5	1	32
4	low	4	5	1	5	5	2	3	2	1	3	31
5	high	3	5	2	3	5	4	5	5	5	3	40
6	low	3	5	1	4	2	3	3	5	4	3	33
7	low	3	5	2	4	1	5	4	3	4	3	34
8	low	5	1	1	4	1	4	3	5	1	2	27
9	high	5	4	2	4	3	3	5	5	4	3	38
10	high	4	3	4	5	3	4	5	5	4	4	41
11	low	3	4	2	1	1	4	5	4	1	1	26
12	low	5	4	4	4	4	3	4	4	5	4	41
13	high	3	2	1	4	5	4	5	5	4	2	35
14	high	1	4	1	5	5	5	3	5	4	3	36
15	low	5	4	2	3	1	1	1	3	2	3	25
16	high	3	5	2	4	5	2	5	4	5	3	38
17	high	1	3	1	2	5	4	4	4	5	3	32
18	low	1	2	5	2	3	3	3	4	5	5	33
19	high	4	5	1	5	5	4	5	5	1	1	36
20	high	3	5	2	4	4	3	5	4	2	1	33
21	high	5	1	3	2	5	4	3	5	4	3	35

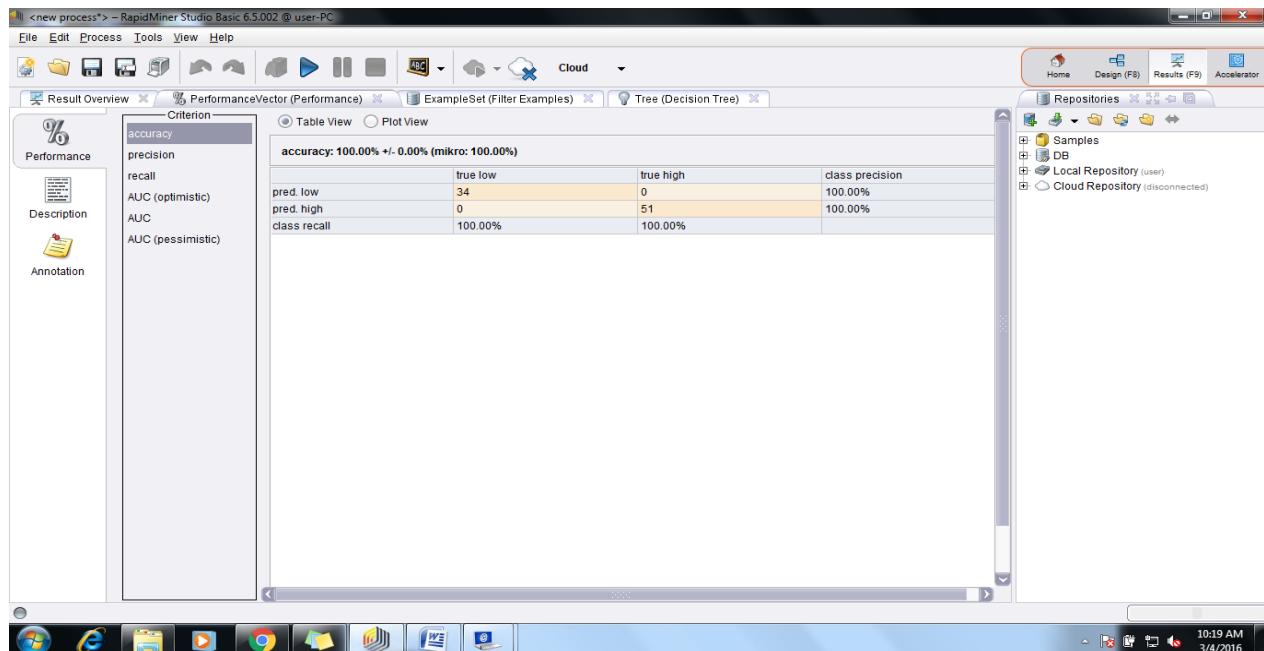
Table: 3

Row No.	LABEL	prediction(L..._confidence(..._confidence(...)	APP	COM	FCLR	FEXP	HBT	MD	RCT	RES	
1	?	low	1	0	5	4	4	4	3	4	4

Table: 4

The label attribute is predicted by the tool with the help of the previously given data set.

If the calculated average value is above the threshold value, then the prediction is ‘high’ and if it is less than the threshold value, the prediction is ‘low’. To be more precise, the predicted value ‘high’ refers to as the people concentrate more on the specified categories and their prediction might not go wrong as well, while the predicted value ‘low’ denotes that they do not concentrate more on the specified categories and so their prediction may not be accurate.

**Table: 5**

Accuracy of this model was determined to be 100%.

## V. CONCLUSION

Various works have been accomplished related to emotions. The main idea of this paper is to predict up to what level a person can observe other behavior when categorized under certain circumstances. The model is an approach that gives a basic idea of how people predict other's behavior and to which category does them emphasis more on to predict it. Further approach can be done by analyzing in depth.

## VI. FUTURE WORK

In this paper, the proposed technique concentrates on the prediction of an individual behavior and the future work here can be extended in the path of colors, where the behavior can be analyzed with the choice of a color.

## REFERENCES

- [1] Predicting Human Behavior via Math ASIAN VOICES IN ORANGE COUNTRY : Sung-Chull Lee, University professor.one in a series of profiles during Asian / Pacific American Heritage Month .May 18,1992|ROSE KIM
- [2] Cognitive Behavioural Systems: Cost 2102 International Training School.edited by Anna Esposito, Antonietta M.Esposito, Alessandro Vinciarelli, Rudiger Hoffmann, Vincent.February 21-26, 2011
- [3] 10 Traits of Emotionally Resilient People Part 1 of 2 in the blog series, "Cultivating Resilience for Total Well-Being". May 21,2013
- [4] EMOTION REGULATION QUESTIONNAIRE (ERQ) Reference: Gross, J.J., & John, O.P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology*, 85, 348-362.
- [5] Scherer, Klaus, and David Sander, eds., 2009. Oxford Companion to Emotion and the Affective Sciences, Oxford: Oxford University Press.
- [6] A Case Study: Use of Applied Behavior Analysis with an Autistic Adolescent(2010) Evan Jaffe Philadelphia College of Osteopathic Medicine.
- [7] BEHAVIOR-ANALYTIC RESEARCH ON DEMENTIA IN OLDER ADULTS(2011) - Maranda A Trahan, SungWoo Kahng, Alyssa B Fisher and Nicole L Hausman.

[8] DATA MINING TECHNIQUES: A SOURCE FOR CONSUMER BEHAVIOR ANALYSIS Abhijit Raorane & R.V.Kulkarni. Department of computer science, Vivekanand College, Tarabai park Kolhapur. January 18<sup>th</sup>,2011

[9] Looking to the Future: Will Behavior Analysis Survive and Prosper?  
Alan Poling(2010).

[10]Workshop Primer: Predicting Human Behavior The Social, Cultural & Ethical Dimensions of “Big Data” March 17, 2014 - New York, NY <http://www.datasociety.net/initiatives/2014-0317/>

## AUTHOR BIOGRAPHY

**D.Radha**, Controller of Examinations and Associate Professor, Department of Computer Science, M.O.P Vaishnav College for Women, Has a teaching experience of 22years and 6 months. She obtained her PG Degree and M.Phil., with second rank in Mathematics from Bharathiyar University. She also holds M.C.A. and M.Phil., degree in Computer Science. She is pursuing Ph.D. in Anna University. Her areas of interest include Data Mining and Discrete Mathematics. Her research work is published in International Journals and she has presented papers in International Conferences. She received Cognizant Technology – Best Faculty Award.



**Sandhya T R**, obatained her UG Degree ‘Bsc.Information Technology’ in Sri Krishna College of Arts & Science, Coimbatore and is currently pursing Msc.Information Technology in the M.O.P Vaishnav College for Women.



**R. Jayaparvathy**, Professor, Department of ECE has a teaching experience of 24 years, 6 months. She graduated with distinction from Government College of Technology, Coimbatore in the year 1987. She obtained her PG degree in Applied Electronics from Coimbatore Institute of Technology, Coimbatore and Ph.D. from AU-KBC Research Centre, Anna University, Chennai. Prior to joining SSN, she served as a Faculty in Coimbatore Institute of Technology, Coimbatore for 24 years. She has also worked as Member-Research Staff in AU-KBC Research Centre for 3 years. She has been the Principal Investigator of AICTE sponsored projects and her prototype project for Elephant Intrusion Detection in the Forest border areas done in collaboration with the Forest Department was widely reported by media. Her areas of interest include Wireless MAC, Wireless Sensor Networks and Embedded Systems. She is a research Supervisor under Anna University, Chennai and Coimbatore.



**Vaishnavi Y**,obatained her UG Degree ‘Bsc.Computer Science’ in Ethiraj College For Womens’s, and is pursing Msc.Information Technology in the M.O.P Vaishnav College for Women.

