Improved e-Governance Model Using Classification Rule Mining and Cluster Analysis

S. Anuradha[1] and E. Kirubakaran[2]

[1](PhD Research Scholar (Full Time), School of Computer Science, Engg. & Applications, Bharathidasan University, Tiruchirappalli, India)

[2](Additional General Manager, Bharat Heavy Electricals Limited, Tiruchirappalli, India)

Abstract: *The higher educational institutions use the information systems for all their academic purposes. e-Governance is used in all the aspects of the higher education sector. Decision making is the key challenge in the current scenario. These institutions have tera bytes of data and using this data for key decision making is the critical issue. If these institutions uses the available data for better planning for their courses, this would be of much useful to the students.This research proposes of use of classification rule mining and clustering analysis data algorithms for improving the e-Governance in higher educational institutions.*

I. INTRODUCTION

India has one of the largest higher education systems in the world. Despite having the largest higher education system, the quality of education, in general, cannot be claimed to be the best. Technical and vocational education in India has seen enormous growth in recent years with a large increase in total number of institutes imparting higher education. On one hand, this growth promises to produce more skilled youth to fulfill needs of ever growing Indian industry and on the other hand it poses a huge challenge for the governing bodies like UGC, AICTE, etc., The introduction ofe-governance in higher education is one such concept that can empower the governing bodies to administer the progress of the education plan in the whole country and serves various stakeholders in a much better way (Dr. R. K. Shrivastava, 2014)

It is important that, the University need to facilitate the students by giving the most current information regarding course selection, educational history and as well as current and future educational and career demands. At present, most of the Universities maintain the database of the student performance for many years and have terabytes of data. Using these information systems, the student can select the courses based on the approved curriculum. Few Universities also implemented e-Advising systems at certain level. These systems lack proper decision making. If data mining is used in this decision-making process, it helps the student in preparing effective educational plans. This is the appropriate time for the Universities to apply data mining in their data and retrieve useful knowledge and share it with all the stake holders (Jiawei Han, 2011).

This paper proposes to conduct research on using a Comprehensive data mining technique, which will support for students in using the E-governance system in a more useful way. Architecture is presented to use the e-Advising in higher educational institutions. The proposed e-Advising system uses software agents for better decision making. The proposed architecture with the developed agents are implemented using a prototype

Objective

The following are the objectives of this research

i. To improve the decision making in the existing e-Governance models in higher education systems with reference to academic monitoring.
ii. Propose architecture by including the appropriate data mining techniques and software agents.
iii. To apply data mining techniques for discovering the hidden relationships in the archive database.

In order to meet the objective, the following steps are considered:

i. The existing e-Governance systems with respect to proposed interactive and improved e-advising system are analyzed in details.
ii. Analyzed the classification rule mining techniques and cluster analsys in the available e-Governance systems.
iii. Identified the appropriate classification and clustering techniques. These techniques are used in the proposed architecture.
iv. A software architecture is proposed to implement the e-Governance model.
v. A prototype is developed by considering the identified data mining techniques, type of software agents, and specialized database. *Microsoft Visual Studio .net platform*is used as front end for prototype development. *MS-Access* database is used as a back end.
vi. The developed prototype is tested to meet the stated objectives.

II. REVIEW OF LITERATURE

E-Advising in National and International Universities

Tony Feghali*et al*(Tony Feghali, 2011) proposes web-based decision support tool (the Online Advisor) that helps advisors and students make better use of an already present university student information system. The author quotes that with this online advisor, advisors can have all the information needed about the courses taken by each student, about the courses still to be taken, about the changes in the academic requirements as they are decided by the appropriate committee(s) on campus, and explore future options for a student. This is said to be done instantly and accurately due to the online advisor's direct connection with the university's student information system. The paper quotes that measures have to be taken to secure that the online advisor delivers what it is designed to do. It has to support the face-to-face meeting between advisor and student and not replace it except for optional prescriptive functions. It has to provide advisors more time to focus on student development and contribute in making better planning and scheduling.

Leora Waldner *et al* (Leora Waldner, E-Advising Excellence: The New Frontier in Faculty Advising, 2011)explores the concept and need for faculty e-advising, defined here as the systematic deployment of online instructional tools in a faculty advising capacity. The author believes that e-Advising brings faculty advising to a medium convenient to online students, and in doing so may improve the quality of advising and student academic success while enhancing online student retention among other benefits. Two basic tenets are discussed, namely, the real-time e-advising and the online organization shell as a resource pod for on- demand access. The paper further explains various factors stimulating e-Advising excellence. The paper highlights few examples of Specific e-Advising strategies using online instructional tools like, Virtual Learning Environment (VLE) Advising Organization, Wimba Pronto Chatroom Office Hours, Classroom Advising Archives, and Orientation Videos. The paper further lists out the limitations of e-advising. The paper concludes that by systematically institutionalizing faculty e-advising, universities can ensure that the future will experience excellence both in online faculty instruction and online faculty advising.

Gary L. Kramer paper (Kramer, 2012) discusses in detail about online e-advising and its role on higher education. Firstly the paper highlights the series of issues that are driving the adoption of online advising. They are, demand for self-service, distance education, timely, convenient access, valuable use of time by advising professionals. The paper then writes about the possible actions for various types of advising. The paper says that many excellent web-based online advising initiatives focus on specific areas, such as course selection, course transfer, or financial planning and few examples have been highlighted in the paper. The paper then discusses the role of e-advising in higher education and the principles of effective online planning. The author concludes that given appropriate planning, web technology and student services providers can form a partnership to deliver high-quality, flexible online advising.

Elizabeth Reed Osika *et al*(Elizabeth Reed Osika, 2007) discusses various ideas about Blackboard course shell to assist in the advising process. The paper describes the academic advising process and suggests utilizing what tools are currently available and repurposing them to assist in the advising process, like the course management system. The author outlined several things that can be done within a blackboard course shell to assist in the advising process. They are reformat the menu, scheduling appointments with the discussion board, progress through a program, communication through private and public means, assessment and satisfaction data, real time notification of changes and other information. The paper also projects the limitation in the use of the course management system, especially the time that is required of the advisor and must be weighed against the benefits realized by the student and the advisor.

Zuhrieh Shana *et al* paper (Zuhrieh Shana, 2014) describes the design and creation of an electronic Student Academic Advising System (SAAS) to support conventional advising procedures. In addition to evaluating the impact of academic e-advising on the student experience in higher education, this study explores the two different types of advising (e-Advising and traditional/in-person academic advising). The Student Academic Advising System (SAAS) is a computer program that gathers all the necessary data about a university student's major such as study plan and student grades, and also collects the personal information of the advisors in a central database to help students plan what individual courses they need to take why and when, which frees up advisers' time to perform higher level advising. The database uses MS-SQL Server database version 2000 and MS Visual Studio tools such as VB.NET and Visual FoxPro. The findings offer awareness into how effective academic e-advising might be used as an alternative/support to traditional, face-to-face advising methods. The author proposes the future plans as: Arabizing the current version of the SAAS, incorporating advisor and student feedback about the SAAS, adding important dates and university events, Addressing unusual cases in advising. The paper concludes with the future challenges as creating interactive academic advising and adding orientation videos to navigate through the e-advising platform.

E-Governance and Data Mining

Dhulipalla Vijay Krishna (Krishna, 2014) proposes role on e-Governance models and its applications to boost e-governance method. e-Governance in an application to supply the government services to the commoner in a very value effective manner. The paper aim at demonstrating the applications of e-governance

models in e- governance applications. This Paper deals with scope and implementation of information deposition; data processing altogether the scale of e-governance like Government to Citizen (G2C), Government to government (G2G), Government to Business (G2B). The analysis work is aimed to represent the potential of information mining within the context of sensible techniques of e-governance. Data processing provides economical techniques for state agencies to investigate information quickly and economically.

Vaibhav Panwar paper (Panwar, 2008) deals with scope and use of data warehousing & Data mining in all the dimensions of e-governance like Government to Citizen (G2C) Citizen to Government (C2G) Government to government (G2G) Government to Business, Government to NGO (G2N). There are many methodology used to increase the efficiency of E-governance. By using techniques Data warehousing, OLAP ,Data mining the paper suggests that data warehousing is very helpful in analyzing Current & Historical data finding useful pattern & support decision strategies, OLAP is useful in solving complex queries & views , interactive online analysis of data.

Dr. Velamala Ranga Rao (RAO, 2014) proposes "A Framework for e-Government Data Mining Applications (eGDMA) – for effective Citizen Services", An Indian Perspective for empowering e-government services in decision making. In this framework, the governments' applications are divided into two, namely; common and department specific applications. These applications are applied and examined through an exhaustive case study and findings and results has been reported. This paper mainly discuss on common & department specific applications of the government organizations and how data mining techniques can help the planners and decision makers in making use of big data for decision making. The framework is proposed to be helpful to the policy makers and planners for extracting the hidden information from various e-Government databases to improve decision making by framing guidelines for new projects for providing better services to the stake holders.

III. RESULTS AND DISCUSSION

Based on the proposed research work architecture, it has been implemented ae-advising recommender system using .Net. This research work uses the following database where all the previous course details are stored as semantic data sources.

Dataset Used

Table 4a: Student Details

S.No	Field Name	Data Type	Size	Description
1	Reg No	Text	Varchar 2(10)	Student Registration Number
2	Specialization	Text	Varchar 2(25)	Subject Specialization
3	Name	Text	Varchar 2(30)	Name of the Student
4	House No	Number	Varchar 2(5)	House Number
5	Street No	Number	Varchar 2(5)	Street Number
6	Sub Street No	Text	Number (10)	Sub Street Number
7	Place	Text	Varchar 2(25)	Name of the Place
8	Country	Text	Varchar 2(25)	Name of the Country
9	E-mail	Text	Varchar 2(30)	Student E-mail Address
10	Telephone	Number	Number (20)	Student Phone Number

Table 4b: Internal and External Assessment

S.No	Field Name	Data Type	Size	Description
1	Semester	Text	Varchar 2(15)	Name of the Semester
2	Year	Number	Number (4)	Academic Year
3	Reg No	Text	Varchar 2(10)	Student Register Number
4	Module Code	Text	Varchar 2(20)	Subject Code
5	Assessment parameter	Text	Varchar 2(30)	Assessment Parameter
6	Max Mark	Number	Number (5)	Maximum Mark for each Subject
7	Marks Scored	Number	Number (5)	Mark Secured
8	Percentage	Number	Number (5)	Total Percentage
9	Group	Number	Number (5)	Student Group

Table 4c: Exam Summary

S.No	Field Name	Data Type	Size	Description
1	Semester	Text	Varchar 2(15)	Name of the Semester
2	Year	Number	Number (4)	Academic Year
3	Module Code	Text	Varchar 2(20)	Subject Code

4	Tutor	Text	Varchar 2(20)	Name of the Tutor
5	Session	Text	Varchar 2(2)	Exam Session
6	Avg_Internal	Number	Number 2(5)	Average Internal Marks
7	Avg_External	Number	Number 2(5)	Average External Marks
8	Avg_Total	Number	Number 2(5)	Average Total
9	Pass_Percentage	Number	Number 2(5)	Student Pass Percentage

Prototype Design

The above database is stored in MS-Access. A prototype is developed to test the framework and implement the algorithms as discussed in Chapters 3 and 4. The user interface is implemented as a web application using ASP.Net and VB.Net. This is a specific prototype for the defined problem and can be generalized with little modifications in the coding part of ASP.Net and VB.Net. The three developed algorithms namely *Apriori, C4.5* and *DBSCAN* is implemented in this prototype.

Results of Classification Tree

The table 4d provides the total 4/5/6 item sets generated. In order to construct a classification tree, the unique item sets from the given sub set are to be identified. Table 5g present the number of unique item sets.

Table 4d: Unique Item Set

Unique Item set		**4**	**5**	**6**
Fall	2007	8	3	0
	2008	8	3	3
	2009	8	5	3
	2010	8	3	0
	2011	5	0	0

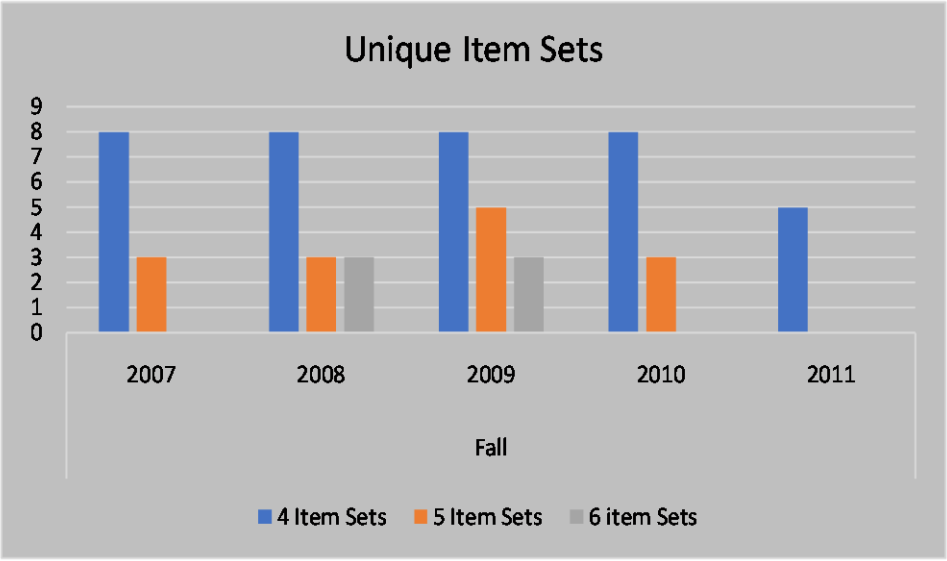

Figure 4a: 6 Unique Item Sets with Minimum Support 60

Interpretation

Figure 4a, depicts the number of unique 4/5/6 item sets generated from the association rule mining technique. This item sets are further considered for classification analysis. The familiar classification tree algorithm, C4.5 is used in this research work to test the discussed framework.

Table 4e presents the first level of classification tree for the four item sets. In the first level, the parent node and the possible child nodes are identified by the C4.5 algorithm. In the first row, the four item set association rule is "IT 0001, IT 0002, CS 0355, CS 038". IT 001 is identified as the root node. IT 002, CS 038, CS 0355 is identified as the child node. For the root node IT 001, the set of all possible child nodes are subsequently provided in the rows from 2 to 5. That is, if a student choses IT 01 as a course, the following are the possible combination that can be considered as the other choices.

a. IT 0002, CS 038, CS 0355
b. IT 0002, CS 044, CS 0355
c. IT 0002 CS 062 CS 035
d. CS 0355 CS 044 CS 038
e. CS 0355 CS 062 CS 038

These combinations are generated in the first level of classification analysis. The second level and third level combinations are not considered due to the fact that the possible generated combination can be easily inferred using the association rule mining result itself

The above justification is applicable to all the rows in table 4e. The interested combination (any row) is considered as input to the next phase. i.e Clustering technique.

Table 4e: Parent Nodes and Child Nodes and Count - 4 itemset

Year	Total Combination	4 Item Set	Parent Count	Parent Node	Child Count	Child Node
2007	**8**	IT 0001, IT 0002, CS 0355, CS 038	1	IT 0001	3	IT 0002, CS 038, CS 0355
		IT 0001, IT 0002, CS 0355, CS 044	1	IT 0001	3	IT 0002, CS 044, CS 0355
		IT 0001, IT 0002, CS 0355, CS 062	1	IT 0001	3	IT 0002 CS 062 CS 0355
		IT 0001, CS 0355, CS 038, CS 044	1	IT 0001	3	CS 0355 CS 044 CS 038
		IT 0001, CS 0355, CS 038, CS 062	1	IT 0001	3	CS 0355 CS 062 CS 038
		IT 0002, CS 0355, CS 038, CS 044	1	IT 0002	3	CS 0355 CS 044 CS 038
		IT 0002, CS 0355, CS 038, CS 062	1	IT 0002	3	CS 0355 CS 062 CS 038
		CS 0355, CS 038, CS 044, CS 062	1	CS 0355	3	CS 038 CS 062 CS 044
2008	**8**	IT 0001, IT 0002, CS 0355, CS 038	1	IT 0001	3	IT 0002 CS 038 CS 0355
		IT 0001, IT 0002, CS 0355, CS 044	1	IT 0001	3	IT 0002 CS 044 CS 0355
		IT 0001, IT 0002, CS 0355, CS 062	1	IT 0001	3	IT 0002 CS 062 CS 0355
		IT 0001, CS 0355, CS 038, CS 044	1	IT 0001	3	CS 0355 CS 044 CS 038
		IT 0001, CS 0355, CS 038, CS 062	1	IT 0001	3	CS 0355 CS 062 CS 038
		IT 0002, CS 0355, CS 038, CS 044	1	IT 0002	3	CS 0355 CS 044 CS 038
		IT 0002, CS 0355, CS 038, CS 062	1	IT 0002	3	CS 0355 CS 062 CS 038
		CS 0355, CS 038, CS 044, CS 062	1	CS 0355	3	CS 038 CS 062 CS 044
2009	**8**	CS 0355, CS 038, CS 044, CS 062	1	CS 0355	3	CS 038 CS 062 CS 044
		IT 0002, CS 0355, CS 038, CS 044	1	IT 0002	3	CS 0355 CS 044 CS 038
		IT 0002, CS 0355, CS 038, CS 062	1	IT 0002	3	CS 0355 CS 062 CS 038
		IT 0001, CS 0355, CS 038, CS 044	1	IT 0001	3	CS 0355 CS 044 CS 038
		IT 0001, CS 0355, CS 038, CS 062	1	IT 0001	3	CS 0355 CS 062 CS 038
		IT 0001, IT 0002, CS 0355, CS 038	1	IT 0001	3	IT 0002 CS 038 CS 0355
		IT 0001, IT 0002, CS 0355, CS 044	1	IT 0001	3	IT 0002 CS 044 CS 0355
		IT 0001, IT 0002, CS 0355, CS 062	1	IT 0001	3	IT 0002 CS 062 CS 0355
2010	**8**	IT 0001, IT 0002, CS 0355, CS 038	1	IT 0001	3	IT 0002, CS 038, CS 0355
		IT 0001, IT 0002, CS 0355, CS 044	1	IT 0001	3	IT 0002, CS 044, CS 0355
		IT 0001, IT 0002, CS 0355, CS 062	1	IT 0001	3	IT 0002 CS 062 CS 0355
		IT 0001, CS 0355, CS 038, CS 044	1	IT 0001	3	CS 0355 CS 044 CS 038
		IT 0001, CS 0355, CS 038, CS 062	1	IT 0001	3	CS 0355 CS 062 CS 038
		IT 0002, CS 0355, CS 038, CS 044	1	IT 0002	3	CS 0355 CS 044 CS 038
		IT 0002, CS 0355, CS 038, CS 062	1	IT 0002	3	CS 0355 CS 062 CS 038
		CS 0355, CS 038, CS 044, CS 062	1	CS 0355	3	CS 038 CS 062 CS 044
2011	**5**	IT 0001, IT 0002, CS 0355, CS 038	1	IT 0001	3	IT 0002, CS 038, CS 0355

		IT 0001, IT 0002, CS 0355, CS 044	1	IT 0001	3	IT 0002, CS 044, CS 0355
		IT 0001, CS 0355, CS 038, CS 062	1	IT 0001	3	CS 0355 CS 062 CS 038
		IT 0002, CS 0355, CS 038, CS 062	1	IT 0002	3	CS 0355 CS 062 CS 038
		CS 0355, CS 038, CS 044, CS 062	1	CS 0355	3	CS 038 CS 062 CS 044

Table 4f presents the first level of classification tree for the four item sets. In the first level, the parent node and the possible child nodes are identified by the C4.5 algorithm. In the first row, the five item set association rule is "IT 0002, CS 0355, CS 038, CS 044, CS 062". IT 002 is identified as the root node CS 0355 CS 038 CS 062 CS044 is identified as the child node. For the root node IT 002, this is the only child node that exists set That is, if student choses IT 02 as a course, the following are the only possible combination that can be considered as the other choices.

"CS 0355 CS 038 CS 062 CS 044"

These combinations are generated in the first level of classification analysis. The second level and third level combinations are not considered due to the fact that the possible generated combination can be easily inferred using the association rule mining result itself.

The above justification is applicable to all the rows in table 4f. The interested combination (any row) is considered as input to the next phase. i.e. Clustering technique. The six item sets are not available for the academic year 2011.

Table 4f: Parent Nodes and Child Nodes and Count - 5 itemset

Year	Total Combination	5 Item Set	Parent Count	Parent Node	Child Count	Child Node
2007	**3**	IT 0002, CS 0355, CS 038, CS 044, CS 062	1	IT 0002	4	CS 0355 CS 038 CS 062 CS 044
		IT 0001, IT 0002, CS 0355, CS 038, CS 044	2	IT0001 CS 038	4	IT 0002 CS 0355 CS 044 CS 038
		IT 0001, IT 0002, CS 0355, CS 038, CS 062	1	IT 0001	4	IT 0002 CS 0355 CS 062 CS 038
2008	3	IT 0002, CS 0355, CS 038, CS 044, CS 062	1	IT 0002	4	CS 0355 CS 038 CS 062 CS 044
		IT 0001, IT 0002, CS 0355, CS 038, CS 044	2	IT0001 CS 038	4	IT 0002 CS 0355 CS 044 CS 038
		IT 0001, IT 0002, CS 0355, CS 038, CS 062	1	IT 0001	4	IT 0002 CS 0355 CS 062 CS 038

2009	5	IT 0001, IT 0002, CS 0355, CS 038, CS 044	2	IT0001 CS 038	4	IT 0002 CS 0355 CS 044 CS 038
		IT 0001, IT 0002, CS 0355, CS 038, CS 062	1	IT 0001	4	IT 0002 CS 0355 CS 062 CS 038
		IT 0002, CS 0355, CS 0362, CS 038, CS 045	1	IT 0002	4	CS 0355 CS 0362 CS 045 CS 038
		IT 0002, CS 0355, CS 0362, CS 038, CS 095	1	IT 0002	4	CS 0355 CS 0362 CS 095 CS 038
		IT 0002, CS 0355, CS 038, CS 044, CS 062	1	IT 0002	4	CS 0355 CS 038 CS 062 CS 044
2010	3	IT 0001, IT 0002, CS 0355, CS 038, CS 044	1	IT 0001	4	IT 0002 CS 0355 CS 044 CS 038
		IT 0002, CS 0355, CS 038, CS 044, CS 062	1	IT 0002	4	CS 0355 CS 038 CS 062 CS 044
		IT 0001, IT 0002, CS 0355, CS 038, CS 062	1	IT 0001	4	IT 0002 CS 0355 CS 062 CS 038
2011	-NIL-					

Table 4g presents the first level of classification tree for the five item sets. In the first level, the parent node and the possible child nodes are identified by the C4.5 algorithm. In the second row, the six-item set association rule is "IT 0002, CS 0355, CS 0362, CS 038, CS 044, and CS 045". IT 002 is identified as the root node CS 0355 CS 0362 CS 038 CS 044 CS 045 is identified as the child node. That is, if a student choses IT 002 as a course, the following is the possible combinations that can be considered as the other choices.

a. CS 0355 CS 0362 CS 038 CS 044 CS 045
b. CS 0355 CS 0362 CS 038 CS 044 CS 091

These combinations are generated in the first level of classification analysis. The second level and third level combinations are not considered due to the fact that the possible generated combination can be easily inferred using the association rule mining result itself.

The above justification is applicable to all the rows in table 4g. The interested combination (any row) is considered as input to the next phase. i.e Clustering technique. The six item sets are not available for the academic year 2007, 2010 and 2011.

Table 4g: Parent Nodes and Child Nodes and Count - 6 itemset

Year	Total Count	6 Item Set	Parent Count	Parent Node	Child Count	Child Node
2007	-NIL-					
2008	3	IT 0002, CS 0355, CS 0362, CS 038, CS 044, CS 045	1	IT 0002	5	CS 0355 CS 0362 CS 038 CS 044 CS 045
		IT 0002, CS 0355, CS 0362, CS 038, CS 044, CS 091	1	IT 0002	5	CS 0355 CS 0362 CS 038 CS 044 CS 091

		CS 0362, CS 038, CS 044, CS 045, CS 045, CS 091	1	CS 0362	5	CS 038 CS 044 CS 045 CS091 CS 045
2009	3	CS 0362, CS 038, CS 044, CS 045, CS 045, CS 091	1	CS 0362	5	CS 038 CS 044 CS 045 CS 091 CS 045
		IT 0002, CS 0355, CS 0362, CS 038, CS 044, CS 045	1	IT 0002	5	CS 0355 CS 0355 CS 038 CS 044 CS 045
		IT 0002, CS 0355, CS 0362, CS 038, CS 044, CS 091	1	IT 0002	5	CS 0355 CS 0362 CS 038 CS 044 CS 091
2010	-NIL-					
2011	-NIL-					

All these results serve as input to the cluster analysis. To proceed further a sample subset (only one row from the table 4e/4f/4g is needed as the input. The subset provided in table 4h is considered in this research work to move to the next phase.

Results of Cluster Analysis

To perform cluster analysis, this research work uses the familiar clustering algorithm DB SCAN.This clustering algorithm considers the input from classification analysis. The following table 4h is being considered to perform cluster analysis using *DBSCAN.*

Table 4h: Input for Cluster Analysis (Minimum Support: 60)

Semester	Year	Parent Node	Child Node
Fall	2009	IT 0001	IT 0002 CS 0355 CS 044 CS 038
		IT 0002	CS 0355 CS 0362 CS 045 CS 038
		IT 0002	CS 0355 CS 0362 CS 045 CS 095
		IT 0002	CS 0355 CS 062 CS 038 CS 044

By considering the row 1 in table 4h as the input, the developed prototype generates the following possible clusters as provided in table 4i.

Table 4i: Cluster Analysis 1

Sessions Handled	Tutor	Average Internal	Average External	Average Total	Pass Percentage
Session-B	Adrian	50	11	61	97
Session-E		40	24	64	90
Session-F		45	15	60	93
Session-G		35	22	57	90
Session-H		41	20	61	90
Session-I		41	10	51	77
Session-K		42	20	62	89
Session-A	Vasanth	34	22	56	85
Session-B		31	22	53	71
Session-C		28	18	46	48
Session-D		35	20	55	80
Session-A		40	17	57	97
Session-B		51	10	61	97
Session-C		45	16	61	97

Session-D	Swamy	35	20	55	80
Session-E		54	10	64	90
Session-F		40	20	60	93
Session-G		37	20	57	90
Session-H		51	10	61	90
Session-I		51	25	76	77
Session-J		64	20	84	96

For the above generated cluster, the attribute '*tutor'* is considered as the primary attribute. This is the output produced by *DBSCAN* algorithm. This table provides the summary of previous courses handled by a tutor. The student can refer this table and decide on course planning.

By considering the row 2 in table 4i as the input, the developed prototype generates the following possible clusters as provided in table 4j.

Table 4j: Cluster Analysis 2

Sessions Handled	Tutor	Average Internal	Average External	Average Total	Pass Percentage
Session-D	Preethika	35	18	53	81
Session-C		33	15	48	59
Session-B		29	19	48	58
Session-A		30	15	46	44
Session-B	Renuka	29	19	48	58
Session-A		30	15	46	44
Session-D	Raja	35	18	53	81
Session-C		33	15	48	59
Session-A	Vasanth	34	22	56	85
Session-B		31	22	53	71
Session-C		28	18	45	48

By considering the row 3 in table 4j as the input, the developed prototype generates the following possible clusters as provided in table 4k.

Table 4k: Cluster Analysis 3

Sessions Handled	Tutor	Average Internal	Average External	Average Total	Pass Percentage
Session-D	Preethika	35	18	53	81
Session-C		33	15	48	59
Session-B		29	19	48	58
Session-A		30	15	46	44
Session-B	Renuka	29	19	48	58
Session-A		30	15	46	44
Session-A	Joseph	31	24	55	79
Session-D	Raja	35	18	53	81
Session-C		33	15	48	59
Session-A	Angeline	31	24	55	79
Session-A	Bharani	30	22	52	80

By considering the row 4 in table 4k as the input, the developed prototype generates the following possible clusters as provided in table 4l.

Table 4l: Cluster Analysis 4

Sessions Handled	Tutor	Average Internal	Average External	Average Total	Pass Percentage
Session-B	Nithyanand	35	15	50	50
Session-B	Chandra	35	15	50	50
Session-D	Preethika	35	18	53	81
Session-C		33	15	48	59
Session-B		29	19	48	58
Session-A		30	15	46	44
Session-B	Renuka	29	19	48	58
Session-A		30	15	46	44
Session-D	Raja	35	18	53	81
Session-C		33	15	48	59
Session-A	Ramu	33	29	62	89

IV. FUTURE DIRECTIONS

The following points can be considered in future to improve the performance of the developed e-Governance model.

i. Incorporation of different data mining algorithms

More data mining techniques and algorithms will be added in the developed model to further enhance the performance of the e-Governance model. At present the prototype uses only the three types of data mining algorithms specified in Chapter 1. The remaining types of data mining algorithms also will be considered to develop a full fledged system prototype.

ii. Data integration

Data needed for data mining process is usually stored in different industrial firms and government agencies using different data formats. Data formats are both structure-specific and vendor-specific. There have been a lot of works on data integration and data exchange, but with little success. These issues have become crucial with the emergence of data warehouses.

iii. Web service

This model is developed as a web service in the future and will be linked with the information systems available with the Universities / Colleges.

iv. Parallel and Distributed Data Mining

Due to the high volumes of spatial data, mining using machines or distributed farms of workstations may improve the performance, especially when dealing with high volume of data. This data can be mined through networks, including internets. New approaches can be necessary for distributed automated data mining.

v. Multidimensional Visualization

Just discovering knowledge is not enough, because it has to be presented in a manner that the user can understand easily. One of the most effective ways of digesting the rules discovered is through graphical visualizations. Humans are very good at interpreting visual data and scenes. This fact should be exploited in the data mining process. Multidimensional data visualization has been studied, but multidimensional rule visualization is still an immature area. Automated data mining can use some well-developed visualization techniques in Computer graphics.

V. REFERENCES

[1] *Dr. R. K. Shrivastava, D. A. (2014). Role of e-Governance to strengthen higher*
[2] *education system in India. IOSR Journal of Research & Method in Education (IOSR-JRME)*
[3] *Elizabeth Reed Osika, A. R. (2007). Online Advising: Stretching Blackboard to Meet Students' Needs. 23rd Annual Conference on Distance Teaching & Learning.*
[4] *Jiawei Han, M. a. (2011). Data Mining: Concepts and Techniques. 3rd edition, Morgan Kaufmann.*
[5] *Kramer, G. L. (2012). Online Advising. EDUCAUSE Centre for Applied Research.*
[6] *Krishna, D. V. (2014). Data Mining in E-Governance Models and Applications. International Journal of Business Policy and Strategy Management , 1(1).*
[7] *Leora Waldner, D. M. (2011). E-Advising Excellence: The New Frontier in Faculty Advising. MERLOT Journal of Online Learning and Teaching, 7(4).*
[8] *Panwar, V. (2008). Role of Data Warehousing and Data Mining in e-Governance.*
[9] *RAO, D. V. (2014, Oct-Dec). A Framework for e-Government Data Mining Applications (eGDMA) for Effective Citizen Services" -An Indian Perspective. International Journal of Computer Science and Information Technology Research, 2(4).*
[10] *Tony Feghali, I. Z. (2011). A Web-based Decision Support Tool for Academic Advising. International Forum of Educational Technology & Society (IFETS), 82-94.*
[11] *Zuhrieh Shana, ,. S. (2014). SAAS: Creation of an e-Advising Tool to Augment Traditional Advising Methods. Computer and Information Science, 7(1).*