A Predictive System for Distance Learning Based on Ontologies and Data Mining

Boufardea Evangelia
University of Patras
CTI "Diophantus" Computer Technology Institute & Press
Patras, Greece
mpoufard@ceid.upatras.gr, mpoufard@westgate.gr

Garofalakis John
University of Patras
CTI "Diophantus" Computer Technology Institute & Press
Patras, Greece
garofala@ceid.upatras.gr, garofala@cti.gr

***Abstract*— The development of distance learning, e-learning and online learning, has increased exponentially in recent years. Lately, researchers have begun to investigate various data mining techniques in order to improve the quality of this type of education. However, although distance learning in education is well established, there are a few attempts to extract educationally useful information during the course and before the final evaluation. In this paper, we propose an ontology based on the structure of a distance learning environment which enriches a recommendation system with rules generated by data mining techniques. Tutors can use this recommendation system in order to predict learners' progress and their final performance. This application will enhance the efficiency of any distance learning or e-learning platform and will be beneficial for learners as well as for tutors in the learning process.**

Keywords: ontology; protégé; RDF; data mining; weka; classification; J48 algorithm; distance learning; HOU

I. INTRODUCTION

The rapid spread of Internet has caused significant changes in many sectors of the economy and society worldwide. From these changes, education could not be left out. With the increasing development of information technologies, a new form of education appears, e-learning (distance education), which revolutionized the educational process.

Furthermore, while the World Wide Web gradually transforms into Semantic Web, new standards and models such as Extensible Markup Language (XML) [20], Resource Description Framework (RDF) and OWL Web Ontology Language are evolving in order to enhance this technology.

The storage, presentation, transmission and search of information according to those standards open up new horizons in the utilization of the Web. Ontologies are increasingly get used for knowledge representation.

A large ontology contains useful data for a system of distance education and thus it is deserved to investigate the "hidden knowledge", i.e., to discover possible associations or to find repeated patterns and forms or extreme events.

This paper proposes a data mining approach to discover relationships between the learning resources metadata. A new recommendation system is developed for assisting tutors in distance learning to predict learners' progress and their final performance. The developed system is based on a new framework using data mining techniques in metadata derived from an ontology. The basic idea was to design an ontology that can store knowledge about the learners' skills in relation to a specific educational purpose. We used for our research PLI23 - Telematics, Internet of the Hellenic Open University, which has a very specific subject and 4 mandatory projects per year, and then we exploited the imported data in order to discover rules and predict learners' progress. The key in distance learning is the communication between learners and tutors. With the help of the proposed system, this communication is getting better, more immediate and effective. In other words, the quality of distance learning is improved.

This paper is organized as follows: in Section II, related work, a quick reference to ontologies, data mining and distance education, which are the basis of our application, are presented. In Section III, we present our implemented application is presented. The application has three parts. The first refers to the ontology that was created in order to represent a course. The second part refers to the data mining techniques that were used in order to exploit the data from the ontology, which algorithms were used and why, which attributes were taken into account, and finally, the rules generated. The third part of the application -the predictive system- is a platform which is used by tutors in order to predict learners' progress. The platform is based on the results of the previous parts. Finally, conclusions and an outline of further work are presented.

II. TOPICS OF INTEREST

Distance learning, ontologies and data mining techniques are the topics that were studied in order to implement our decision support system. There is related work that is presented below.

A. *Distance learning*

Higher education systems all over the world are challenged nowadays by the new information and communication technologies (ICT). Distance education or distance learning is a field of education that focuses on teaching methods and technology with the aim of delivering teaching, often on an individual basis, to students who are not physically present in a traditional educational setting such as a classroom. It has been described as "a process to

 ISBN: 978-1-61208-218-9

create and provide access to learning when the source of information and the learners are separated by time and distance, or both." Distance education courses that require a physical on-site presence for any reason (including taking examinations) have been referred to as hybrid or blended courses of study [18].

Distance learning programs can fundamentally change the way schools compete for students, especially part-time students. A school that develops distance learning programs usually increases the scale and scope of its offerings. Many traditional institutions have added distance learning programs. Academic institutions and corporations are combining resources to bring distance learning programs to workplaces. Academic institutions offer courses through distance learning so students have opportunities to create a degree program that uses course offerings from multiple schools.

Hellenic Open University (HOU) is the 19th Greek State University, but the only one that provides distance education in both undergraduate and postgraduate levels via the development and utilization of appropriate learning material and methods of teaching. The HOU's mission is to provide distance education at both undergraduate and postgraduate level. For that purpose, it develops and implements appropriate learning material and methods of teaching. The promotion of scientific research as well as the development of the relevant technology and methodology in the area of distance learning falls within the scope of the HOU's objectives [22].

B. Ontology

Ontologies are widely used in Knowledge Engineering, Artificial Intelligence and Computer Science, in applications related to knowledge management, natural language processing, e-commerce, intelligent integration information, information retrieval, database design and integration, bio-informatics, education, and in new emerging fields like the Semantic Web [1].

The term is borrowed from philosophy, where ontology is a systematic account of existence.

“Ontology is an explicit specification of a conceptualization.” This definition became the most quoted in literature and by the ontology community. Based on Gruber’s definition, many definitions of what ontology is were proposed. Borst (1997, page 12) modified slightly Gruber’s definition as follows: Ontologies are defined as a formal specification of a shared conceptualization.” [7].

Several technologies have been developed for constructing and developing the Semantic Web. RDF and its extensions such as OWL have been developed to define metadata schemas, domain ontologies and resource descriptions. RDF is a W3C (World Wide Web Consortium) standard developed in 1997. It is a standard model for data interchange on the Web. RDF has features that facilitate data merging even if the underlying schemas differ, and it specifically supports the evolution of schemas over time without requiring all the data consumers to be changed [23]. Web ontology language (OWL) built on RDF is the new W3C recommendation for ontology construction with facilitates for effective reasoning capabilities by consistency checking through inference rules such as transitivity, symmetry etc. OWL is designed for use by applications that need to process the content of information instead of just presenting information to humans. OWL facilitates greater machine interpretability of Web content than that supported by XML, RDF, and RDF Schema (RDF-S) by providing additional vocabulary along with a formal semantics. OWL has three increasingly-expressive sublanguages: OWL Lite, OWL DL, and OWL Full [24].

Eyharabide et al. [13] described an ontology that was implemented for predicting students' emotions when interacting with a quiz about Java programming. Zhuhadar et al. [9] and Pin-Yu Pan et al. [12] proposed a hybrid recommendation strategy of content-based and knowledge-based, in which aiming to filter recommended items from the available items according to the user's preferences. Sridharan et al. [3] presents a multi-level ontology - driven topic mapping approach to facilitate an effective visualization, classification and global authoring of learning resources in e-learning.

In this paper, we implemented an application ontology that contains all the definitions needed to model the knowledge required for the particular area of application. Application ontologies often extend and specialize the vocabulary of the domain for a given application such as ours. For instance in this work, we created an ontology for a system specialized in distance learning for the HOU and we adapt the particularities of HOU into the ontology.

C. Data mining techniques

Data mining involves the use of sophisticated data analysis tools to discover previously unknown, valid patterns and relationships in large data sets. These tools can include statistical models, mathematical algorithms, and machine learning methods (algorithms that improve their performance automatically through experience, such as neural networks or decision trees). Consequently, data mining consists of more than collecting and managing data, it also includes analysis and prediction [8].

Data mining is ready for application in the business community because it is supported by three technologies that are now sufficiently mature [6]:

- Massive data collection
- Powerful multiprocessor computers
- Data mining algorithms

The most commonly used techniques in data mining are: Artificial neural networks (non - linear statistical data modeling tools, usually used to model complex relationships between inputs and outputs or to find patterns in data) [15], Decision trees (a decision support tool that uses a tree-like graph or model of decisions and their possible consequences, including chance event outcomes, resource costs, and utility) [17], Genetic algorithms (a search heuristic that mimics the process of natural evolution) [19], Association Rules (popular and well researched method for discovering interesting relations between variables in large databases) [16], etc.

There are several approaches to data mining techniques in e-learning generating recommendations based on a user's profile. Burdescu et al. [4] proposed a novel structure of a support system for e-Learning infrastructure that is based on data representing learner's activities and processes. Markov chain modeling and classification is used as main intelligent procedure for data analysis. Hanna [10] described how we can profit from the integration of data mining and the e-learning technology. Taking into consideration the suggested methods, data mining can be used to extract knowledge from e-learning systems through the analysis of the information available in the form of data generated by their users. In this case, the main objective is to find the patterns of system usage by teachers and students and, perhaps most importantly, to discover the students' learning behavior patterns. In this paper, we used classification process and more specific Decision Trees in order to find a set of models that describes and distinguishes data classes and concepts. The derived model is based on the analysis of a set of training data.

A dissertation also with subject "Understanding dropout of adult learners in e-learning" [2] has been also conducted, which is related, too.

III. APPLICATION

A. Building the Ontology

Nowadays, ontologies offer the ability to model the knowledge of a domain in a discrete and definite way. We used Protégé as a framework application. Figure 1 shows the design of "pli" ontology in Protégé.

As far as concerning HOU, by means of an ontology, it will be possible to describe the knowledge domain the subjects constituting it, the relations among the various subjects, as well as methodologies and means with which they are presented. The content of an ontology depends both on the amount of information and on the degree of formality that is used to express it. Generally, two main types of ontologies are distinguished: lightweight and heavyweight. In this paper, the lightweight approach is adopted according to this definition of ontology: "ontology may take a variety of forms, but it will necessarily include a vocabulary of terms and some specification of their meaning". A lightweight ontology is a structured representation of knowledge, a taxonomy where the concepts are arranged in a hierarchy with a simple relationship between them.

In this ontology, we had to model and represent the relevant aspects and domains of knowledge for a distance learning environment which contains the knowledge about the followings:

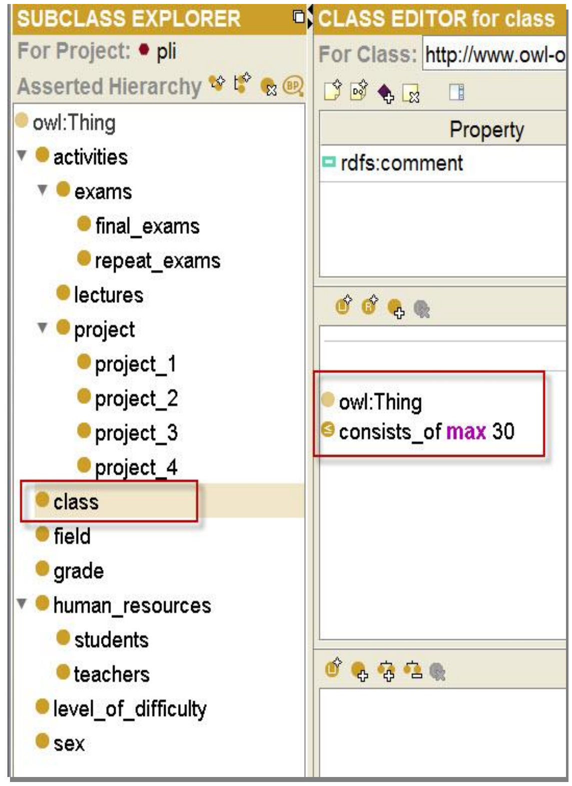

Figure 1. Pli ontology

Apart from the above classes, there are also relations, datatypes and restrictions such as following in Table I:

TABLE I. ELEMENTS OF PLI ONTOLOGY

Pli ontology	
Relations	"Students belong to a class" or "A class consists of students" "A project has level of difficulty" "Teachers teach at a class" or "A class is taught by teachers" "Students take activities" or "Activities are taken by students" "Students are of sex" "Students have grades" etc…
Datatypes	Exams_type (**string**) – domain: *exams, students* Grade (**float**) – domain: *students, grade* Pass_exams (**Boolean**) – domain: *students*
Restrictions	A class <= consists of max 30 (students), etc…

In our study, there were three different fields, three classes, four teachers, seventy nine students, and three level of difficulty for the projects. In order to check if there is any

 ISBN: 978-1-61208-218-9

inconsistency of our ontology we used a reasoner. Pellet [14] reasoner was used, which is a piece of software able to infer logical consequences from a set of asserted facts or axioms.

After the ontology building, metadata are gathered, and the title and description fields of the metadata XML/RDF files are separated. The next step is to use the produced RDF in order to take the data in an ARFF (Attribute-Relation File Format) format in order to process them with data mining techniques. An ARFF file is an ASCII text file that describes a list of instances sharing a set of attributes.

B. Data mining techniques for prediction

Data-mining techniques have been applied in order to find interesting patterns, build descriptive and predictive models from large volumes of data accumulated through the ontology. The results of data mining can be used for getting a better understanding of the underlying educational processes especially in distance learning, for generating recommendations and advices to students, for improving resource management, etc.

We used classification as it has many applications in both traditional education and modern educational technology. The best results are achieved when classifiers can be learned from real data, but in educational domain the data sets are often too small for accurate learning. Our sample was only 79 instances.

Our most important concern was to select a sufficiently powerful model, which catches the dependencies between the class attribute and other attributes, but which is sufficiently simple to avoid overfitting. Both data preprocessing and the selected classification method affect this goal. We used Weka for the extraction of rules that predict students' progress. Weka is a collection of machine learning algorithms for data mining tasks. The algorithms can either be applied directly to a dataset and Weka [21] contains tools for data pre-processing, classification, regression, clustering, association rules, and visualization. It is also well-suited for developing new machine learning schemes. It can also make comparative surveys or checks in datasets. Someone can see the errors of every classification with a popup menu, or if there is a decision tree the result can be shown as a picture.

The produced RDF from the ontology gives us the instances of the attributes that we are interested in. The attributes that are important for our predictive system are the grades of the projects, the grades of the exams, the type of exam, the gender of the student. Some other important attributes could be the age of the students, their marital status, their educational background etc. Unfortunately, due to the privacy of the personal data we could not have access to these data so we restricted in the referred attributes.

Real data is often incomplete in several ways. It can contain missing or erroneous attribute values. Erroneous values are generally called noise, but in some contexts noise means only measurement errors. Educational data is usually quite clean because it is either collected automatically (log data) or checked carefully (students' scores). In our system we used real data based on students' grades so we avoid the noise.

All the useful needed data are in the ontology that was implemented previously, so we can extract them in an arff format, see Figure 2, via RDF schema, with the help of a parser. We need these data in arff format as it is the suitable format for processing of data through Weka. For this reason firstly we codified students' data in a tuple in order to process them if there are hidden relations among them i.e., students whose 2nd and 3rd project had grade up to 7,5 succeeded in final exams, or students who did not submit the 1st project they dropped out the course.

The following data were elaborated:

Sex (male/female)	Grade 1st project (number)	Grade 2nd project (number)	Grade 3rd project (number)	Grade 4th project (number)
Grade Final exams (number)	Grade Repeat exams (number)	Type of exams (final/repeat)	Pass (true/false)	

For example, a student can be represented as tuple at Weka as follows:

male, 5, 6, 7, -1, 4, 9, repeat, true

That is, the student is man and his grades at the first three projects are 5, 6, 7, he did not submit the fourth project, he failed at final exams with grade 4 and he gave repeat exams with grade 9 and he passed the course.

```
@relation pli

@attribute student {female,male}
@attribute grade1 real
@attribute grade2 real
@attribute grade3 real
@attribute grade4 real
@attribute gradeexam real
@attribute graderepeat real
@attribute typeofexam {final,repeat}
@attribute pass {true,false}

@data
male,9.3,7.9,7.5,6.0,7.2,-1.0,final,true
female,7.9,7.8,6.2,9.0,5.7,-1.0,final,true
male,9.1,8.8,9.6,9.4,8.1,-1.0,final,true
```

Figure 2. Arff file for pli

The dataset that was created consists of 79 instances and every record has 9 attributes. Essentially it is investigated whether a student will pass the exams of the course based on rating criteria the grades of the projects and the grades of exams.

Through the process of data, Weka extracts rules applicable to a proportion of cases (confidence). We are interested in those that apply for 100% of cases. The data classification is a supervised learning process in which an apprentice algorithm takes a number of observations (records) as a basis for its training.

As we use the dataset in Weka, we can have immediately a visualization of all attributes depending on pass or fail status of student.

The data mining technique that was used in our application was classification. Decision trees are maybe the best-known classification paradigm. A decision tree represents a set of classification rules in a tree form. Decisions tree are a collection of nodes, branches, and leaves. Each node represents an attribute; this node is then split into branches and leaves. Decision trees work on the "divide and conquer" approach; each node is divided, using purity information criteria, until the data are classified to meet a stopping condition.

The earliest decision trees were constructed by human experts, but nowadays they are usually learned from data. One of the best known algorithms is C4.5 [11] is an algorithm used to generate a decision tree. The basic idea in all learning algorithms is to partition the attribute space until some termination criterion is reached in each leaf. Usually, the criterion is that all points in the leaf belong to one class. However, if the data contains inconsistencies, this is not possible. As a solution, the most common class among the data points in the leaf is selected. An alternative is to report the class probabilities according to relative frequencies in the node. J48 is an open source Java implementation of the C4.5 algorithm in the Weka data mining tool.

For our recommendation system we decided to use J48 algorithm since decision trees have many advantages: they are simple and easy to understand, they can handle mixed variables (i.e., both numeric and categorical variables), they can classify new examples quickly, and they are flexible. Enlargements of decision trees can easily handle small noise and missing attribute values. Decision trees have high representative power, because they can approximate nonlinear class boundaries, even if the boundaries are everywhere piecewise parallel to attribute axes [5]. The idea of classification is to place an object into one class or category, based on its other characteristics

We run our dataset using J48 algorithm and cross validation 10 folds. This means that the dataset is divided into 10 parts. The nine of them used in order to train the algorithm and the rest one is applied to trained algorithm. We made many tests giving as data only the grade of 1st project and the grade of final and repeat exams, and then giving the grade of 1st and 2nd project etc. We will present the decision tree which is resulted from the data of all the grades of projects and the sex of students. In Figure 3 you can see this decision tree. The building of the tree is shown in the Figure 4.

From the 79 instances the 63 classified correctly and 16 classified erroneously. The precision of the classification of our model is 85.2%.

- Correctly Classified Instances 63 79,7%
- Incorrectly Classified Instances 16 20,3 %

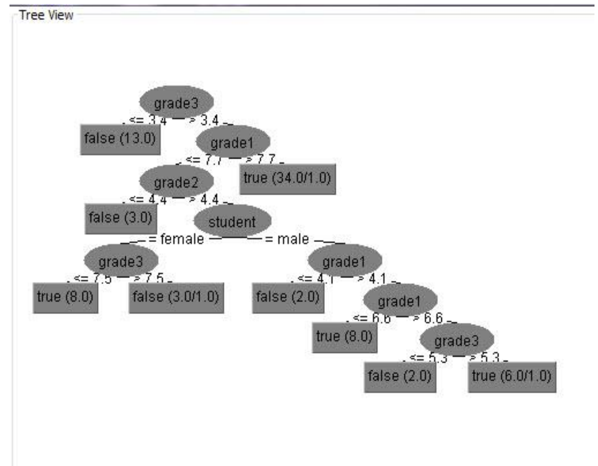

Figure 3. Decision tree

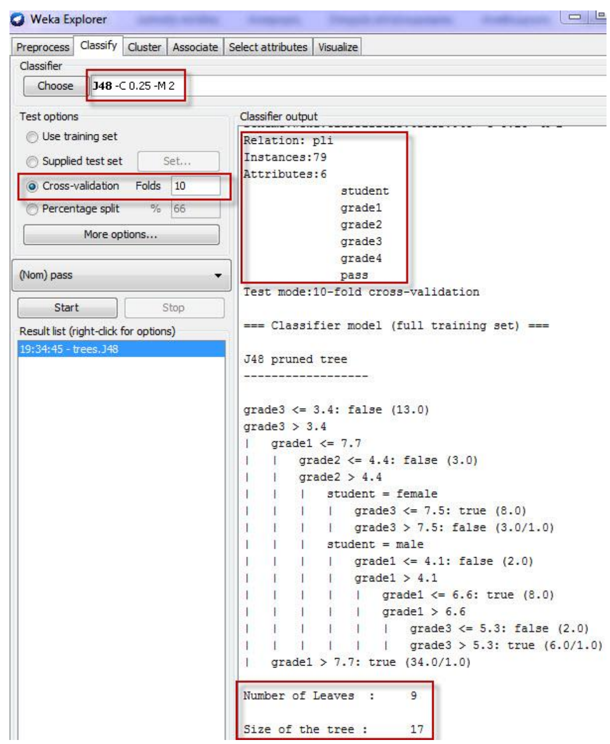

Figure 4. Building the tree with J48 (cross-validation 10 folds)

We observe that for the class "pass=true" (the course) is True Positive=0,891, False Positive=0.417, Precision=0.831, instead of the class "pass=false" is True Positive =0.583, False Positive =0.109 and Precision=0.7.

According to Confusion Matrix the 10 b instances from 24 b (b: false) classified as a (a: true) although there were b so FP=10/24=0.417. On the other hand 6 a classified as b so FP for the "pass=false" is 6/55=0,109.

 ISBN: 978-1-61208-218-9

- Probabilities of being correct given that your decision.
  -Precision of pass=true is 49/59 = 83%
- Probability of correctly identifying class.
  -Recall accuracy for pass=true is 49/55 = 89%
- Accuracy: # right/total =63/79 =~79,7%

We used the same dataset to train J48 algorithm with parameter percentage split 66%. The building of the tree is shown in the Figure 5.

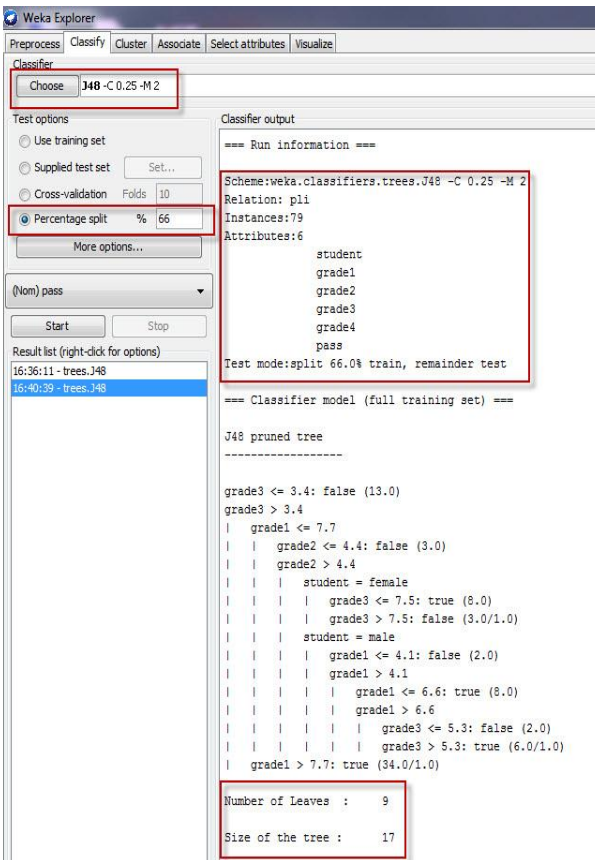

Figure 5. Building the tree with J48 (percentage split 66%)

The metrics are the following: from the 27 instances the 23 instances classified correctly and 4 classified erroneously. The accuracy of our model is 85.2%

- Correctly Classified Instances 23 85.2%
- Incorrectly Classified Instances 4 14.8 %

For the class "pass=true" is True Positive=0.952, False Positive=0.5, Precision=0.87. For the class "pass=false" is True Positive =0.5, False Positive =0.048 and Precision=0.75. According to the Confusion Matrix 3 b instances out of 6b (b: false) classified as a (a: true) although there were b so FP=3/6=0.5. On the other hand 1 a classified as b and for "false" was FP=1/21=0.048.

- Probabilities of being correct given that your decision.
  -Precision of pass=true is 20/23 = 87%
- Probability of correctly identifying class.
  -Recall accuracy for pass=true is 20/21 = 95.23%
- Accuracy: # right/total =23/27 =~85.2%

Summarizing for the case of class "pass=true" the metrics for the two different parameters are shown in Table II. For the class "pass=false", the metrics are shown in Table III. We take into account only the grades of projects and students' sex. As we see the accuracy is the same independently of the parameter and the decision tree is the same too.

TABLE II. METRICS FOR PASS=TRUE

Pass=true	TP	FP	Precision	Recall	Accuracy
J48 (cross-validation 10 folds)	0,891	0,417	0,831	0,891	79,7%
J48 (percentage split 66%)	0,952	0,5	0,87	0,952	85,2%

TABLE III. METRICS FOR PASS=FALSE

Pass=false	TP	FP	Precision	Recall	Accuracy
J48 (cross-validation 10 folds)	0,583	0,109	0,7	0,583	79,7%
J48 (percentage split 66%)	0,5	0,048	0,75	0,5	85,2%

Based on the produced decision trees some rules were created. We used these rules in order to implement a recommendation system.

C. Recommendation system

The goal of the recommendation system is to provide a modern web based environment that will allow the timely assessment of the achievement of learning outcomes, thus supporting the realization of efficient personalized learning paths. The aims are to support the monitoring of learners' progress and provide indicators of successfully achieving learning outcomes, to support the assessment of learning paths and to provide a means for interaction between learners and tutors.

This predictive system is a web based platform that records learners' progress and provide tools for analyzing their performance and estimating the chances of finally achieving the planned learning outcomes based on the rules that generated through the data mining techniques. This platform will be targeted to learners, tutors of HOU.

This study examines the background information from data that impact upon the study outcome of PLI23-

 ISBN: 978-1-61208-218-9

Telematics students at the HOU. Classifying students based on grades information and the rules presented for each node would allow the tutors to identify students who would be "at risk" of dropping the course very soon in academic year. Then the tutor can support student with additional educational material, and give them orientation, advising, and mentoring programs, that could be used to positively impact the academic successes of such students.

Our application is a web based application implemented with open source technologies such as php, MySql, javascript, html. Tutors and learners have access to the platform, but tutors have access in an additional functionality of platform. This functionality is the prediction of students' progress. Tutors can log on in the platform and insert student' data such as name, grades, etc. then they can check with the given data the probability of failing or succeeding in the exams. If the tutors see that a student is "at risk" can communicate with him/her and advise him/her, give supportive material or just make a warning. So the learner could improve his/her performance, study hardier etc.

Figure 6 shows the prediction of "pass" of a student. The tutors can see the percentage of probability and if they want to see they can also see the decision tree that resulted in this percentage. Of course if we had access in more information about students such as age, occupation, marital status etc this system would give us more interesting rules. But as we already referred we could not have access to such information. However, if there are institutes or organizations which don't face problems related to privacy, this system could cover their needs too.

The system also creates arff files automatically after the insertion of data, which means that creates a larger dataset that can be exploited by Weka and maybe it can also change the rules that have been already generated.

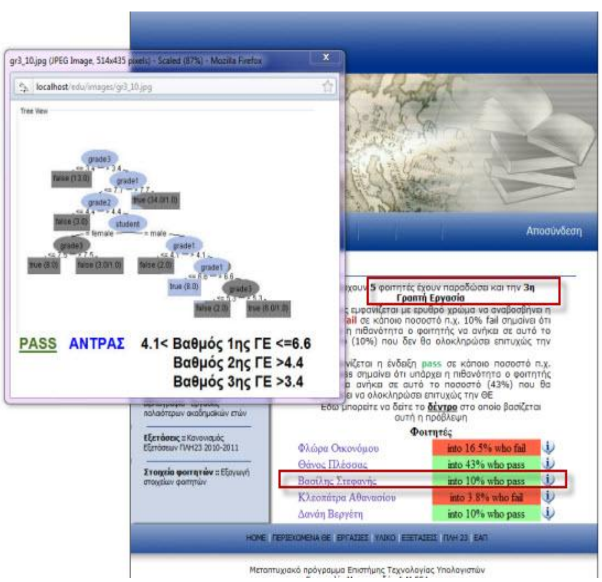

Figure 6. Prediction of a student's final performance

The key in distance learning, as we already mentioned, is the quality of communication. A system like ours can improve the quality of this type of education since it supports the immediate communication between tutor and learner. Tutor has only to click on the name of student and send him/her an email or make a call and inform him/her for his/her progress.

We made piloting testing with real data of students of next years and the results were very satisfying. We use 31 records of students. The application firstly gave the proportion based on the grades of 1st project, then based on grades of 1st and 2nd project and finally based on grades of 1st, 2nd and 3rd project. The grade of 4th project doesn't play any role to the generated rules from previous step of data mining. This happens because of the curricula of HOU, the 4th project is optional. If a student doesn't have good grades in previous projects can give the 4th project in order to be able to participate in final exams. That means that the majority of students do not make the 4th project. The application seems to be very accurately as it has small deviation error i.e., for the 1st project and for the 1st and 2nd project the system predicts that all users will pass the exams. The real data show that only two students fail out from 31. So the system has accuracy 93.5%. Taking into account and the 3rd project the precision of system is getting lower. The system predicts erroneously that 2 students will succeed in exams and 2 students will fail. So the accuracy of system is 87%. However, it predicts that 2 students (women) will marginally fail but at the end they pass. So we assume that this path of the decision tree is weak and if we had a larger dataset to train the algorithm, this path would be different.

The Precision of this pilot dataset is 90.3% and the Recall is 96.6% comparing to the precision and the recall of the classifier performance we see that is more accurately. This happens as the pilot test dataset is too small and the data are too clear. Table IV shows briefly the results of pilot testing.

TABLE IV. PREDICTION OF STUDENTS' FINAL PERFORMANCE

Results/ Projects	Prediction of pass	Prediction of fail	Pass	Fail	Total
1st	31	0	29	2	
1st & 2nd	31	0	29	2	**31**
1st & 2nd & 3rd	29	2	27	4	

IV. CONCLUSIONS AND FUTURE WORK

This application can be the vehicle for improving the quality of existing distance -and not only- learning systems and it will be very helpful not only for the learners who will be able to be informed for their progress but also for the tutors who can improve their teaching process. After the study of different data mining techniques, we chose the one that was most suitable for our application and as we see is the most widespread in this field. The results were very satisfying and after pilot testing of our application we saw that the final results verified the generated rules with small deviation error prediction. The basic idea of this system can be also applied in other distance learning or e-learning systems.

A comparison between some classification tree models would be conducted in the future in order to determine the best model for the dataset. Maybe an alternative to a classification tree should be considered in order to compare the results.

This research is based on background information only. Leaving out other important factors as previous studies, number of courses completed motivation, financial aids, age, marital status, etc.) that may affect study outcome, could distort results obtained with classification trees. For example, including the educational background of students after the submission of the first project would probably improve predictive accuracy of the model. To improve the model, more attributes could be included to obtain prediction with fewer errors.

Furthermore, an idea is to incorporate the C4.5 algorithm to our platform in order to run again and again the algorithm with the new data too. So the model will be always up to dated and new rules could be generated. These rules could be also incorporated to our ontology as restrictions. This enrichment of ontology may lead to the extraction of prediction directly from the ontology.

REFERENCES

[1] A. Gomez-Pιrez, M. Fernαndez-Lopez and O. Corcho, “Ontological Engineering with examples from the areas of Knowledge Management, e-Commerce and the Semantic Web”, First Edition

[2] B. A. Jusung Jun, “Understanding dropout of adult learners in e-learning”, A Dissertation Submitted to the Graduate Faculty of the University of Georgia in Partial

[3] B. Sridharan, H. Deng, B. Corbitt, “An ontology-driven topic mapping approach to multi-level management of e-learning resources”, 17th European Conference on Information Systems, pp. 1187-1198

[4] Burdescu, D.Dan Mihaescu, M. Cristian Ionascu, C. Marian Logofatu, Bogdan, “Support system for e-Learning environment based on learning activities and processes” (19-21 May 2010), Research Challenges in Information Science (RCIS), 2010 Fourth International Conference, pp. 37-42

[5] C. Romero, S. Ventura, M. Pechenizkiy, R.S.J.d. Baker, “Handbook of Educational Data Mining, Edited”, chapter 5

[6] F. Castro, A. Vellido, À. Nebot, F. Mugica, “Applying Data Mining Techniques to e-Learning Problems”, Evolution of Teaching and Learning Paradigms in Intelligent Environment In Evolution of Teaching and Learning Paradigms in Intelligent Environment, Vol. 62 (2007), pp. 183-221

[7] Gruber, Thomas R. ,“A translation approach to portable ontology specifications”, (June 1993), Knowledge Acquisition 5, pp.199–220.

[8] J. W. Seifert, “Data Mining: An Overview”, National Security Issues, pp. 201-217

[9] L. Zhuhadar, O. Nasraoui, R.Wyatt and E. Romero, “Multi-model Ontology-based Hybrid Recommender System in E-learning Domain”, 2009 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology – Workshops, pp.91-95

[10] M. Hanna, (2004) "Data mining in the e-learning domain", Campus-Wide Information Systems, Vol. 21 Iss: 1, pp.29 - 34

[11] N. V. Chawla, “C4.5 and Imbalanced Data sets: Investigating the effect of sampling method, probabilistic estimate, and decision tree structure”, ICML’03 Workshop on Class Imbalances

[12] Pin-Yu Pan, Chi-Hsuan Wang, Gwo-Jiun Horng, Sheng-Tzong Cheng, “The development of an Ontology-Based Adaptive Personalized Recommender System”, Electronics and Information Engineering (ICEIE), 2010 International Conference On, 1-3 Aug. 2010, pp. V1-76 - V1-80

[13] V. Eyharabide, A. Amandi, M. Courgeon, C.Clavel, C. Zakaria, J. Martin, “An ontology for predicting students' emotions during a quiz. Comparison with self-reported emotions”, Affective Computational Intelligence (WACI), 2011 IEEE Workshop on, 11-15 April 2011, pp. 1-8

[14] http://clarkparsia.com/pellet/protege/ [retrieved: March, 2012]

[15] http://en.wikipedia.org/wiki/Artificial_neural_network [retrieved: May, 2012]

[16] http://en.wikipedia.org/wiki/Association_rule_learning [retrieved: May, 2012]

[17] http://en.wikipedia.org/wiki/Decision_tree [retrieved: May, 2012]

[18] http://en.wikipedia.org/wiki/Distance_education [retrieved: March, 2012]

[19] http://en.wikipedia.org/wiki/Genetic_algorithm [retrieved: May, 2012]

[20] http://en.wikipedia.org/wiki/XML [retrieved: May, 2012]

[21] http://www.cs.waikato.ac.nz/ml/weka [retrieved: March, 2012]

[22] http://www.eadtu.eu/hellenic-open-university-hou-greece.html [retrieved: March, 2012]

[23] http://www.w3.org/RDF/ [retrieved: May, 2012]

[24] http://www.w3.org/TR/owl-features/ [retrieved: May, 2012]

 ISBN: 978-1-61208-218-9